Success and failure in cultural markets

Matthew J. Salganik

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2007

ABSTRACT

Success and failure in cultural markets

Matthew J. Salganik

This dissertation is 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. Here we propose 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 novel experimental design we found support for our ideas in a series of four experiments involving a total of 27,267 participants.

Contents

List of Tables

List of Figures

Acknowledgments

It seems strange that this dissertation has only one author given the number of people who have helped me in such important ways. Academically, I owe a huge debt to my adviser Duncan Watts who made this dissertation possible and guided me every step of the way. Also, I would like to thank Peter Dodds for many helpful and encouraging conversations. Peter Hausel built the musiclab website with design help from Jason Booher-Jennings. After the site was built, Alison Binkowski and Robert Stewart kept it running. I also received lots of advice and support from the members of the Collective Dynamics Group and many of the faculty and students of the sociology department.

This work was also made possible by generous funding that I received from Cornell University, Columbia University, the Institute for Social and Economic Research and Policy, the Netherland-America Foundation, and the National Science Foundation. This financial support gave me the luxury of dedicating my time and attention to my research.

Finally, my friends and family were always there to listen to my dissertation problems, and more importantly, to help put them in perspective. Special thanks go to my brother and parents for all their support. And finally, there is Amanda, to whom I can never say thank you enough.

To Amanda

Chapter 1

Introduction

1.1 Success in cultural markets

This dissertation is 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.

Examples of extreme inequality abound. Michael Jackson's *Thriller* has sold more than 41 million copies, while thousands of bands struggle to even release an album (Vogel, 2004, Ch. 5); the first six books in the *Harry Potter* series have collectively sold more than 300 million copies at a time when most authors only achieve sales in the thousands (Wapshott, 2005); the 1997 blockbuster movie *Titanic* grossed $600 million at the U.S. box office alone, nearly fifty times the average U.S. box office take for movies produced in that year[1]; Pablo Picasso's painting *Boy with a Pipe* sold for $104 million dollars at a 2004 auction (Velthuis, 2005, p. 197), a

[1]All box office figures come from `www.boxofficemojo.com`, a source which is routinely cited by the *Los Angeles Times* and *The Wall Street Journal*.

price that was more than 10,000 times the average yearly art-earnings of American artists (Robinson and Montgomery, 2000).

Although each of these cases are unique, they are all part of a generic pattern in cultural markets; most entrants do poorly while a few successes stand out. This pattern leads cultural markets to be called "superstar" (Rosen, 1981) or "winner-take-all" (Frank and Cook, 1995) markets. This tremendous variation in success occurs in many different "popular culture" markets such as pop music (Chung and Cox, 1994; Strobl and Tucker, 2000; Fox and Kochanowski, 2004; Krueger, 2005), trade publishing[2] (Hirsch, 1972; Vogel, 2004; Sorensen, 2007), television (Bielby and Bielby, 1994), and movies (Faulkner and Anderson, 1987; De Vany and Walls, 1997, 2004; De Vany, 2004). This same inequality also occurs in "high culture" markets such as classical music[3] (Gilmore, 1993; Dowd et al., 2002) and visual art (Baumol, 1986; Peterson, 1997; Velthuis, 2005).

Such large differences in success seem to suggest that these superstar objects are somehow inherently different from their peers. Yet, here we encounter the second-half of the puzzle: experts are unable to predict which objects will become superstars. Again examples abound. The first book in the *Harry Potter* series, which went on to sell tens of millions copies, was rejected by eight publishers before finally being accepted for publication (Lawless, 2005).[4] The mega-hit television show *American*

[2]Five authors, John Grisham, Tom Clancey, Danielle Steel, Michael Crichton, and Stephen King, accounted for 70% of total fiction sales in 1994 (Greco, 1997).

[3]Between 1954 and 1969 about 30% of all performances by the top U.S. symphony orchestras were written by just five composers: Beethoven, Mozart, Brahms, Wagner, and Tchaikovsky (Dowd et al., 2002). However, concentration at the top has actually gone down over time. From 1842-1857 the top five of the era (Mendelssohn, Beethoven, Weber, Mozart, and Spohr) accounted for 52.3% of all performances by top U.S orchestras. (Dowd et al., 2002).

[4]Interestingly, Nigel Newton, the chief executive of Bloomsbury and the man who purchased the rights to the first Harry Potter book, attributes the discovery of Harry Potter not to any forethought on his part, but to his daughter who read the manuscript and liked it; she was able to convince him to pay a mere £2,500 for the rights to publish the book. On the phenomenal success of the book and its sequels, the publisher said "We hit it lucky" (Lawless, 2005).

Idol, which at the time of this dissertation is in its sixth season, was initially rejected by three major networks.[5] In music, Bob Dylan's "Like A Rolling Stone," which was recently named the greatest rock and roll song of all time by two major pop music magazines, was almost never released because executives thought it was too long (Considine, 2004).

The incorrect predictions are not limited to missing things that go on to become superstars; sometimes products that are predicted to become superstars fail. For example, Carley Hennessy was a young Irish pop singer whose demo recording greatly impressed the president of MCA Records. Expecting her to become a superstar, MCA invested more than $2 million in the production and marketing of her first album, "Ultimate High." Despite this support, the album sold just 378 copies in its first three months (Ordonez, 2002). Stories of such "flops" are often unknown to those outside of cultural industries who see only successful products. A few other notable flops include the musician Ben Kweller (Seabrook, 2000, Ch. 4), the book *The Interpretation of Murder* (Trachtenberg, 2006), and movies such as *Cleopatra* (Parish, 2006) and *Heaven's Gate* (Caves, 2000, Ch. 8). Those outside of cultural industries are often surprised to learn that failures greatly outnumber the successes. For example, it is estimated that about 80% of movies and that 90% musicians lose money for their parent companies (De Vany and Walls, 2004; Ordonez, 2002; Vogel, 2004).

Not surprisingly, this unpredictability is well known to industry executives. In music, executives routinely talk about the unpredictability of success (Peterson and Berger, 1971; Hirsch, 1972; Denisoff, 1975), and even successful ones refer to their work as "a crap game" (Denisoff, 1975, p. 93). In the film business one successful producer

[5]The story of how *American Idol* eventually made it onto the air bears some surprising parallels to the story of Harry Potter's publication. The key person in getting *American Idol* on the air was not the show's agent, producer, or star, but rather Liz Murdoch—daughter of Rupert Murdoch, the owner of Fox Networks. After she told her father that she liked the British version of the show, *Pop Idol*, Mr. Murdoch ordered his executives to sign and air the show despite their hesitations (Carter, 2006, Ch. 9).

described the work as "rather like being a wildcat oil driller. You hit a lot of dry wells hoping for a gusher" (Faulkner and Anderson, 1987, p. 886), and the well-respected screenwriter William Goldman wrote that, "the single most important fact, perhaps, of the entire movie industry is that 'nobody knows anything'." (Goldman, 1983). Similar sentiments are heard from both television and publishing executives (Gitlin, 1983; Hill and Power, 2005). For example, Leslie Moonves, Chairman of CBS said, "You never really do know what will work or not work. Most of your big hits come out of nowhere" (Hirschberg, 2005) and according to Judith Regan, president of Regan Books, the same rule applies to publishing: "You know, nobody knows" (Williams, 2005).[6]

These impressions of industry executives about unpredictability are also supported by outside research. A study of over 2,000 movies found that the budget of a movie, the name recognition of its stars, its genre, and most other obvious variables, explained relatively little of the variation in success, leading the authors to conclude that in film, "revenue forecasts have zero precision" (De Vany and Walls, 1999). Similar results were found in a study of the success of television shows by Bielby and Bielby (1994).

It is this paradox of cultural markets—that success is at once unequal and unpredictable—that we investigate here. First, we review the literature in this area and propose a conceptual model which resolves the apparent paradox. More specifically, we propose that social influence, a process well studied at the individual level, leads to unanticipated consequences at the collective level including inequality and unpredictability of success. Later in the dissertation we will describe an experimental test of our conceptual model in which we created a web-based "music market" where

[6]Despite, or perhaps because of, the well known unpredictability, some industry executives seem to believe that there exists a magic fomula which will solve their problems. In the 1940's Geroge Gallup brought the then new science of political pooling to Hollywood. His predictions met with limited success (Ohmer, 2006, Ch. 6). More recently, others have sold more elaborate technological approaches (Gladwell, 2006). We suspect that they will met with limited success as well.

27,267 participants listened to and downloaded new pop music.

1.2 A proposed solution

Both the unpredictability and inequality of success for cultural objects have been investigated previously, but the two features are often considered separately. Questions relating to the unpredictability of success have been considered mostly by sociologists working in the production of culture school (Peterson and Anand, 2004). This work has tended to take unpredictability as a given, and then examine its effect on firms or individuals. For example, some work has examined how organizational structure in cultural industries is affected by unpredictability (Peterson and Berger, 1971; Hirsch, 1972; Faulkner and Anderson, 1987), while other work has focused on rhetorical framing strategies used by decision-makers in these highly uncertain environments (Bielby and Bielby, 1994). This body of research goes a long way to understanding the consequences of unpredictability, but since it takes unpredictability as given, it does little to illuminate its causes.

Research on the inequality of success, on the other hand, has mostly been done by economists who have created mathematical models which reproduce this inequality (Rosen, 1981; Adler, 1985; MacDonald, 1988; Chung and Cox, 1994; De Vany and Walls, 1996; De Vany and Lee, 2001) and measured the inequality of success in different markets and at different times (Chung and Cox, 1994; Strobl and Tucker, 2000; Fox and Kochanowski, 2004; De Vany and Walls, 2004; De Vany, 2004; Krueger, 2005). However, Rosen's original work, which has come to dominate the field, completely ignores the unpredictability in cultural markets. Further, the models that do deal with the unpredictability (Adler (1985), De Vany and Walls (1996), De Vany and Lee (2001)) are impossible to test empirically because they require comparisons of multiple realizations of the same stochastic process.

This dissertation is intended to complement these two separate strands of research by proposing that both inequality and unpredictability can arise from the same source: social influence. Here we define social influence to be phenomena whereby the behavior of an individual is affected by the behavior of others. The existence of social influence on individual decisions has been well established by sociologists and experimental social psychologists since the 1930's (Sherif, 1937; Asch, 1952; Katz and Lazarsfeld, 1955) and there has been a substantial body of research attempt to understand the possible *origins* of social influence; for reviews see Bond and Smith (1996); Cialdini (2001); Cialdini and Goldstein (2004). However, there has been little research on the *consequences* of social influence on collective outcomes and this is where we will focus our attention. Next, we will argue that social influence likely operates on decisions in cultural markets, and then we will show that this individual level influence process leads to both inequality and unpredictability of collective outcomes.

1.2.1 Social influence

In cultural markets individuals make two main types of decisions. First, before an individual can "choose" a cultural object,[7] they must first decide which object to even consider. This decision can be difficult because individuals are faced with so many objects; for example, a typical "big-box" bookstore currently stocks between 40,000 and 100,000 books, and those shopping on the Internet see their choices increased by a factor of 10 (Brynjolfsson, Hu, and Smith, 2003). Out of a huge mass of objects, and with no information about most of them, an individual must pick one to consider. In making this decision, individuals are probably more likely to consider objects that are already successful (Adler, 1985; Hedström, 1998; Goldstein and Gigerenzer, 2002; Bonabeau, 2004; Laureti et al., 2005). This mimicry could derive

[7]The choice of the word "choose" is meant to apply to objects that most people can acquire, for example books, and to objects that most people cannot acquire, for example French Impressionist paintings.

from individuals using the behavior of others as a signal of the quality of the objects or because individuals desire to conform to the behavior of other, either in the quest for discussion partners or out of a desire for social approval.[8]

Second, once an individual has decided which object to consider, they must then decide whether to "choose" the object or not. As with the first decision about which object to consider, this second decision is also likely influenced by the behavior of others. Individuals may be more likely to "choose" popular objects, and reject unpopular objects, in order to gain social approval. Additionally, the opinions of others could directly affect the individual's evaluation of the object. For example, individuals laugh longer and more often at jokes in the presence of canned laughter (Fuller and Sheehy-Skeffington, 1974), and individuals evaluate a given text more positively when it is attributed to a high-status author than when it is attributed to a low-status author (Willer, 2005).

These individual tendencies to mimic the behavior of others at both decision steps are often reinforced by structural features of many cultural markets. For example, booksellers often give best-selling books more prominent in-store placement, making them more likely to be considered, and sell them at lower prices, making them more likely to be purchased once they are being considered. Similar patterns appear in music where some radio stations desire to play whatever music is "hot," thereby influencing individuals at both decision steps. Thus, even without knowing the popularity of the objects, individuals can be affected via this indirect influence.

[8]Going back to the study of Deutsch and Gerard (1955), researchers have drawn a conceptual distinction informational and normative motives for conformity. However, in practice it is difficult to distinguish between the two (Kaplan and Miller, 1987; Cialdini and Goldstein, 2004), even when using sophisticated brain imaging techniques (Berns et al., 2005). Therefore, in this dissertation, we will generally lump these two processes together, with the exception of Chapter 6 where we will attempt to distinguish them.

1.2.2 Cumulative advantage

As the above discussion suggests, the social influence process can depend on a combination of individual and societal factors. Although these details may matter for some aspects of aggregate outcomes (Dodds and Watts, 2004, 2005; Bruch and Mare, 2006), we argue that the issues of interest here—unpredictability and inequality—depend only on the most general features of social influence. Because many of the influence processes presented so far have the effect of making successful objects more successful, cultural markets can be understood with the help of a large body of research on what has come to be called "cumulative advantage" models.

Sociologists are probably most familiar with cumulative advantage from research on inequality (DiPrete and Eirich, 2006), status hierarchies (Gould, 2002), and the sociology of science (Merton, 1968; Allison et al., 1982). However, formal models incorporating some version of cumulative advantage have been studied in many fields under a variety of names: rich-get-richer, success-breeds-success, preferential attachment, Gibrat's principle, constant or increasing returns to scale, and the Matthew effect[9] (Newman, 2003; DiPrete and Eirich, 2006). The literature on cumulative advantage can be roughly divided into two distinct classes, one dealing with the inequality of outcomes for a set of objects and the other dealing with the unpredictability of outcomes for specific objects.

The first group of literature, mostly produced by statisticians and physicists, is concerned with inequality in the distribution of outcomes for a set of objects. For example, these researchers have presented models attempting to reproduce distributions of firm sizes, distributions of citations to scientific articles, and distributions of links to specific web-pages (Newman, 2003). Work in this stream of literature,

[9]The name "the Matthew effect" was proposed in Merton (1968) and comes from a passage in the New Testament, gospel of Matthew. In the King James translation, both Matthew 13:12 and 25:29 read, "For unto everyone that hath shall be given, and he shall have in abundance; but for him that hath not shall be taken away, even that which he hath."

beginning with Yule (1925), and later Simon (1955), de Solla Price (1976), Barabási and Albert (1999), Krapivsky and Redner (2001), and many others, suggests that cumulative advantage processes generate inequality in outcomes for objects that are otherwise identical. These dynamic models, while differing in their specifics, share the feature that objects which have become successful tend to become more successful. This process leads to a type of snowballing where initially small differences—possibly the result of random fluctuations—grow into large differences. Taken together, this body of work suggests that the observed inequality in cultural markets is consistent with a cumulative advantage process.

The second part of the cumulative advantage literature, mostly produced by economists, focuses on the outcomes for specific objects instead of the distribution of outcomes. The prototypical example of this work is that on competing technology standards—for example the QWERTY vs. Dvorak keyboards (David, 1985). Research in this literature, for example Arthur (1989), emphasizes that in situations where cumulative advantage operates, it is not possible to predict ahead of time which technology standard will become dominant. Because of "path dependence" and "lock-in," one realization of the process will produce one winner, but the next realization, performed under identical conditions, will produce a different winner. Models of this kind, therefore, suggest that if social influence leads to cumulative advantage, then the outcomes in cultural markets will *never* be precisely predictable because these outcomes vary from realization to realization (Watts, 2002, 2003).[10] Thus, the inability of industry executives to predict successful objects may not stem from a lack of competence, but rather from the inherent impossibility of the task.

[10]An example of a familiar system that has inherent uncertainty is rolling a die. Because the outcome of the die varies across realizations, no amount of information about the die, or previous outcomes, can allow for precise prediction of future outcomes.

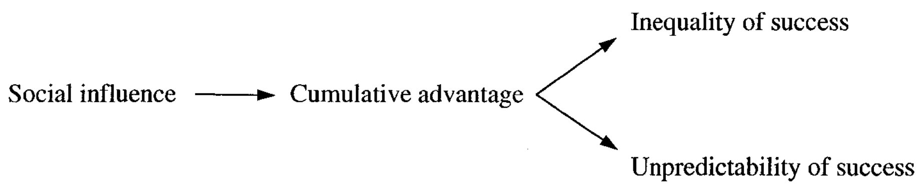


Figure 1.1: Schematic of the conceptual model. The inequality and unpredictability of success, which seem to be contradictory at the aggregate level, can both arise from social influence at the level of the individual leading to cumulative advantage in the success of objects.

1.3 Testing the model

The conceptual model presented above, and shown in figure 1.1, is grounded in existing literature and suggests a solution to the paradoxical nature of success in cultural markets: the inequality and unpredictability in cultural markets, which superficially seem to be at odds with one another, could both stem from social influence at the level of the individual. Unfortunately, this prediction is difficult to test with observational data. To test the claim that social influence leads to inequality of success one must observe indistinguishable groups of people evaluating the same set of objects where in one group individuals make decisions in the presence of social influence (interdependently) and in the other group individuals make decisions in the absence of social influence (independently). Further, to test the claim that social influence leads to inherent unpredictability—that is, truly a crap game—one would need to observe multiple "histories" or "realizations" of the social influence condition. In the same way that one needs to see multiple rolls of dice to see if the outcomes are possible to predict, one needs to observe multiple histories to determine the extent to which the outcomes for cultural objects are inherently uncertain. Without observing these counterfactual conditions, any tests of cumulative advantage models with observational data will be open to multiple, conflicting interpretations (Sobel, 1996; Winship and Morgan, 1999).

To resolve these methodological issues, we have adopted an experimental approach.[11] Specifically, we examined the effect of social influence on the success of cultural objects by creating a controlled "market" for previously unknown pop music. Within our experimental framework we controlled the information that participants had about the behavior of others, and therefore clearly and directly tested our prediction that social influence leads to both inequality and unpredictability of success.[12]

1.4 Dissertation outline

The bulk of the dissertation describes a series of four related web-based experiments summarized in table 1.1.[13] These all explore different aspects of the proposed model and offer support in a variety of ways. The remainder of the dissertation will be organized into the following chapters.

First, **Chapter 2** will describe the general experimental design used, as well as its particular instantiation in the MusicLab website. **Chapter 3** will present the results from these first two experiments which were similar except for the amount of social influence we attempted to impose on the participants. Based on the conceptual model discussed previously, we had two specific, falsifiable predictions. First, we expected that in both of the experiments, the success of the songs in the social influence condition would be both more unequal and more unpredictable than success

[11]Experimental approaches, while not common in sociology, have been used previously. Examples include Berger et al. (1977); Ridgeway (1982); Cook et al. (1983); Lawler and Yoon (1996); Lovaglia et al. (1998); Willer (1999); Molm et al. (2000); Simpson and Macy (2004).

[12]The experimental approach also insures that we will have complete data on *all* the songs in our population; whereas observational studies of cultural objects almost always involve some form of sampling bias because more successful objects are more likely to be included. Following White and White (1965) in their study of French painting, we note that studying a complete population of objects is often a more appropriate way to learn about cultural markets.

[13]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.

	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 1.1: 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.

of songs in the independent condition. Second, because participants in experiment 2 were subject to stronger form of social influence, we expected that success in the social influence condition in experiment 2 would be more unequal and more unpredictable than success in the social influence condition in experiment 1. Showing this dose-response relationship, whereby we are able to manipulate an individual level process and yield predictable results at the aggregate level would be strong support for our conceptual model. Some of the results from this chapter have been published in a paper that was co-authored with Peter Dodds and Duncan Watts (Salganik et al., 2006).

Participants from experiments 1 and 2 were mostly drawn from `www.bolt.com`, a website popular with American teenagers. However, this leaves open the possibility that the experimental findings might only be applicable in that specific population (Park and Lessig, 1977; Sears, 1986; Pasupathi, 1999; Peterson, 2001). **Chapter 4** will address this concern by describing the results of experiment 3 which was identical to experiment 2 except the subject pool was recruited via emails to participants in an unrelated experiment (Dodds et al., 2003).

In experiments 1, 2, and 3, we allowed the popularity of the songs to emerge naturally, without any intervention. However, in real cultural markets, firms and artists actively intervene in this process by marketing and promoting their products, often attempting to distort the perceived popularity of objects. A natural way to think about the effects of these attempts is in the context of the work on self-fulfilling prophecies developed by Robert K. Merton.[14] As defined by Merton (1948), "[a] self-fulfilling prophecy is, in the beginning, a *false* definition of the situation evoking a new behavior which makes the originally false conception come *true*. [emphasis in original]" **Chapter 5** presents the results from experiments 3 and 4 which explore the possibility of self-fulfilling prophecies by providing participants with a false sense of the popularity of the songs.

Chapter 6 probes the data from the four experiments to answer a number of questions that fall between the cracks of the previous chapters. For example, in this chapter we will attempt to determine the relative magnitude of the social influence on subjects' listening, rating, and download decisions. Further, in this chapter we will explore individual heterogeneity in behavior. Finally, **Chapter 7** will address potential future research.

---

[14]Merton based his work on earlier work by W. I. Thomas who developed what Merton called the Thomas theorem which states, "If men define situations as real, they are real in their consequences." (Merton, 1995)

Chapter 2

Experimental design and website

2.1 Abstract design

As stated in Chapter 1, this dissertation will argue that social influence at the individual level leads to both inequality and unpredictability of success at the collective level. To test this idea, one would need to observe collective outcomes both in the presence and absence of social influence. Therefore, the four experiments presented here use the design presented in figure 2.1. 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

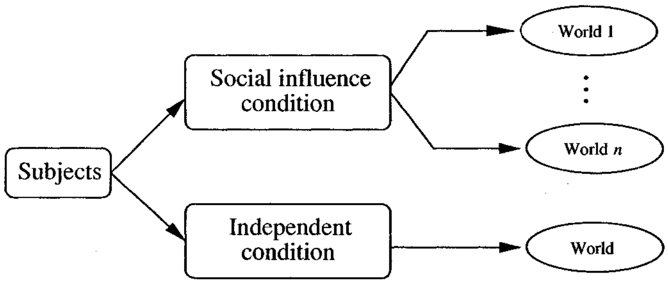


Figure 2.1: Schematic of the experimental design. This design has two main features. First, it allows researchers to isolate the difference in aggregate outcomes when social influence is present or absent. Second, the design allows researchers to observe multiple realizations of the same process, and thus understand the role of chance in collective outcomes.

thought of as multiple, parallel "histories" or "realizations."[1] Because of random assignment, both conditions had indistinguishable groups participants evaluating the same songs. Any differences in aggregate outcomes, therefore, can only be explained by the presence of social influence at the individual level. This sociological research design differs from those common in economics and psychology because rather than focusing on the origins of social influence at the individual level, it focuses on how the social influence process aggregates to produce collective outcomes (Schelling, 1978; Coleman, 1990; Hedström, 2005).[2]

One practical problem with this experimental design is that it requires a very large number of participants. Experiments in psychology and economics often require *hundreds* of participants because they use the individual as the unit of analysis. How-

[1]We only need one independent realization because the behaviors of the participants in this condition are independent by definition. Therefore, as the sample size got the large the behaviors of the groups would converge.

[2]These group level experiments have some precedents in social psychology, especially in the subfield of jury behavior. An incomplete list of examples includes: Sherif (1937), Burleson et al. (1984), Diamond (1997), Bornstein (1999), Schkade et al. (2000), Horowitz and Bordens (2002). Another closely related experimental design was used in Chase et al. (2002) in which groups of fish repeatedly formed dominance hierarchies. Unlike human experiments, however, the fish hierarchies were formed repeatedly for exactly the same group of fish, whereas studies with humans often compare results across similar, but not identical, groups.

ever, the experiments presented here use collective outcomes as the unit of analysis, and since we needed several hundred participants for each collective outcome, this experimental design required *thousands* of participants in total. Ultimately, the four experiments presented here involved a total of 27,267 participants, a number that would be impossible to accommodate in traditional laboratory-based experiments. Fortunately, new technology and widespread Internet access allowed us to conduct web-based experiments, thus avoiding the logistical constraints of physical laboratory space.[3]

Aside from this practical advantage of running web-based experiments, there were two main substantive advantages. First, with our web-based experiments, we were able to reach a much wider group of participants than would have been possible in a standard study of university students in the United States. For example, about 30 percent of our participants lived outside the U.S. and about 65 percent were outside the age-range of traditional university students (table 2.1). This more diverse participant pool means that our results are more likely to reflect general regularities in human behavior rather than tendencies specific to university students in the U.S. (Sears, 1986; Pasupathi, 1999; Peterson, 2001).[4] In addition to allowing us to collect a large and diverse (albeit non-random) sample, the second major benefit of conducting web-based experiments was that they allowed us to observe participants making decisions in a natural environment, not a laboratory. Because participants were not in a laboratory, however, we had much less control over participant recruitment and behavior, creating the possibility of experimental contamination. We took a number of steps, described in section 2.5, to account for these problems.

[3]More information about web-based experiments can be found in Gosling et al. (2004), Birnbaum (2004), Kraut et al. (2004), and Skitka and Sargis (2005).

[4]While our sample was diverse, it was not random. Because participants were not selected with known probability of selection from a sampling frame, our sample should *not* be used to make descriptive statements about any population. However, our experimental results are not based on such statements. Instead, we are comparing the behavior of groups of sample members in different conditions.

Category	www.bolt.com Experiment 1 ($n = 7,149$) (% of participants)	Experiment 2 ($n = 7,192$) (% of participants)	Small-world experiment 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 2.1: 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 (Dodds, Muhamad, and Watts, 2003) and the subsequent web postings these emails generated. Participants in experiments 3 and 4 were older and more international.

2.2 Participant experience during the experiment

Before being able to listen to and download the music, subject had to complete a number of preliminary activities. 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 2.4. The use of 48 songs—the 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 (Caves, 2000; Vogel, 2004).

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. For reasons that will be described later, 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. Further, participants were not told that all their actions during the experiment—listening, rating, and downloading—were logged in our database along with time-stamps for each action. The time-stamps allow us to measure how long each participant spent deciding which song to listen to, how long they listened to each song, and how long

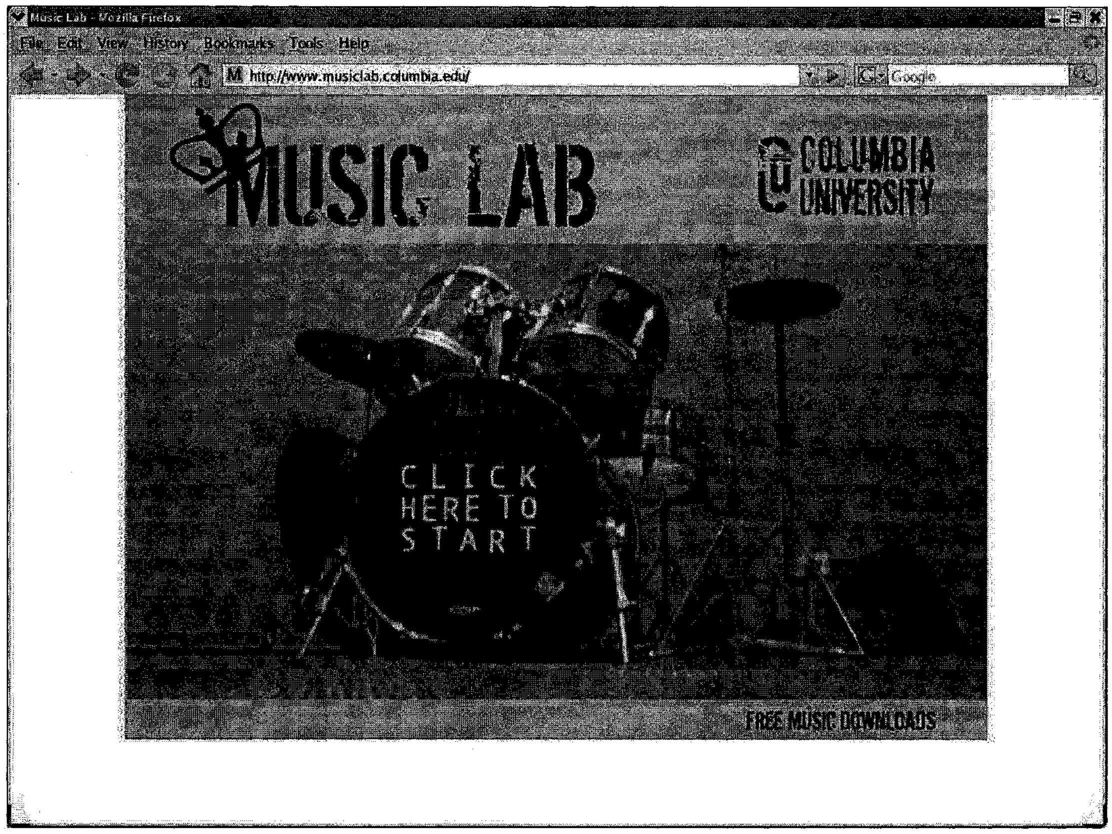


Figure 2.2: Splash screen from the website.

they spent deciding whether to download the song.

Screenshots from all steps of the experiment and presented in figures 2.2 to 2.11. The consent forms, screen text, and survey questions are presented in appendix A.

2.3 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

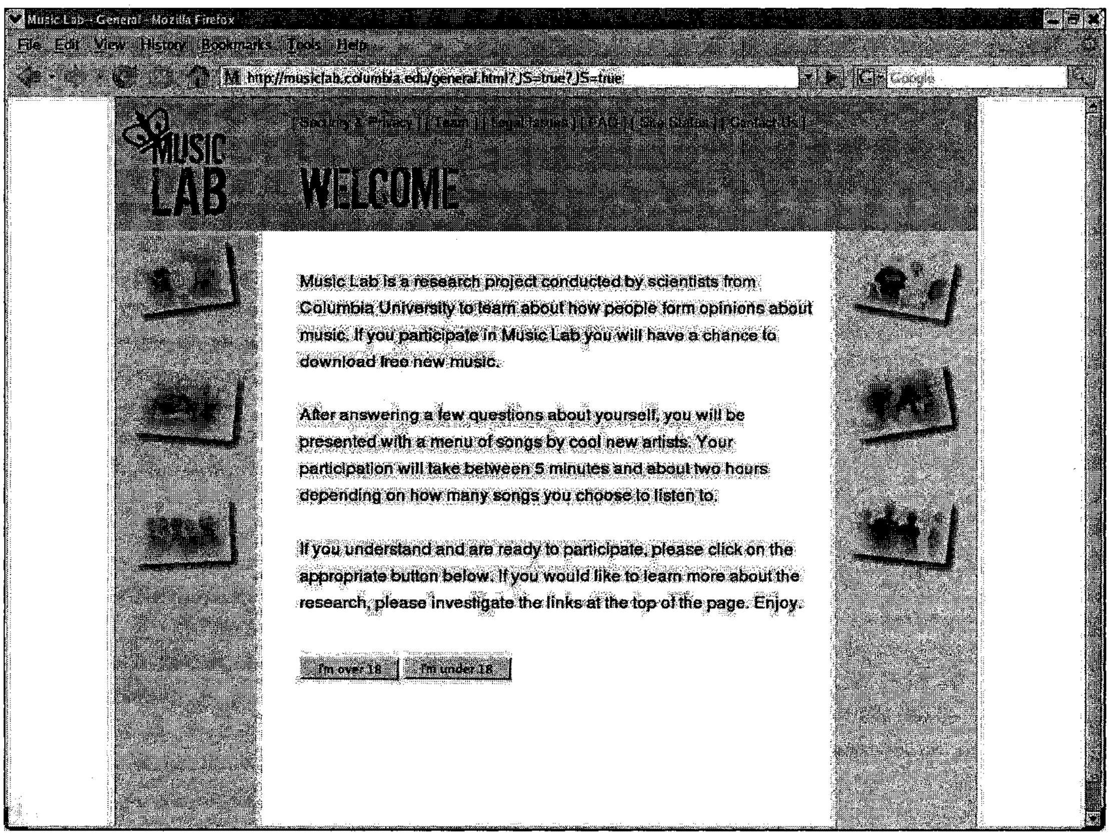


Figure 2.3: Welcome screen from the website.

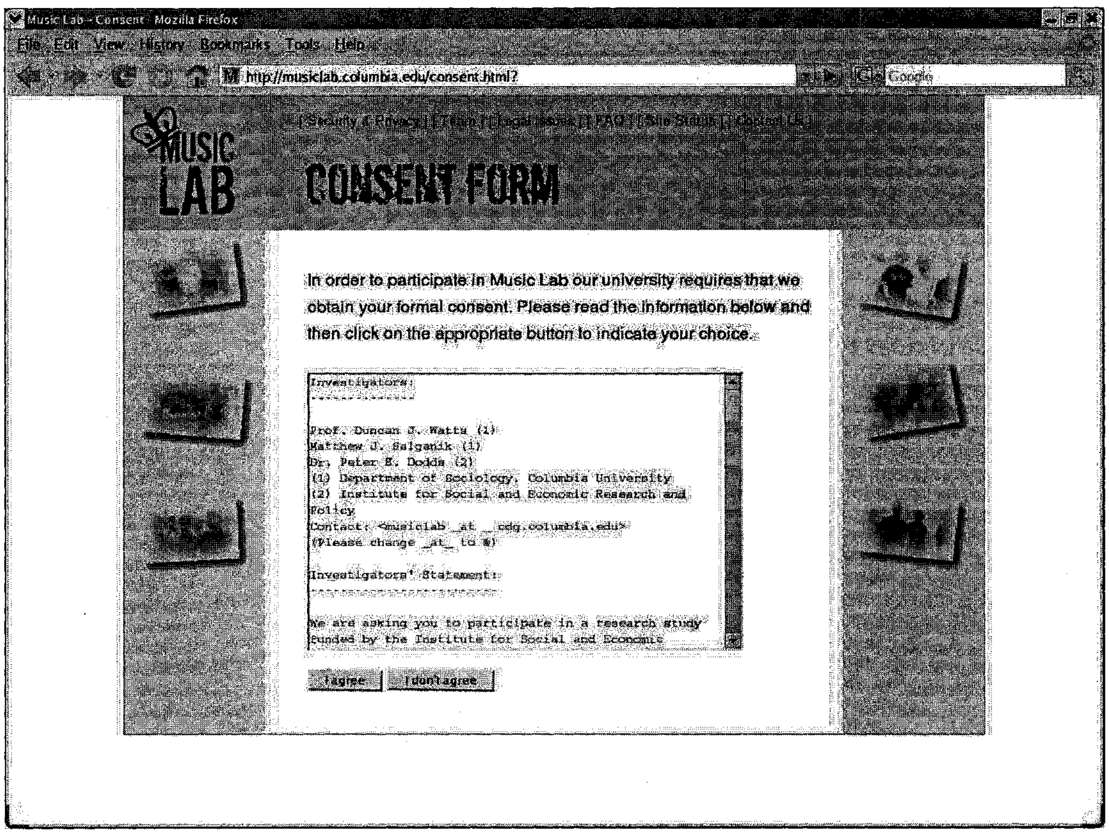


Figure 2.4: Consent screen from the website.

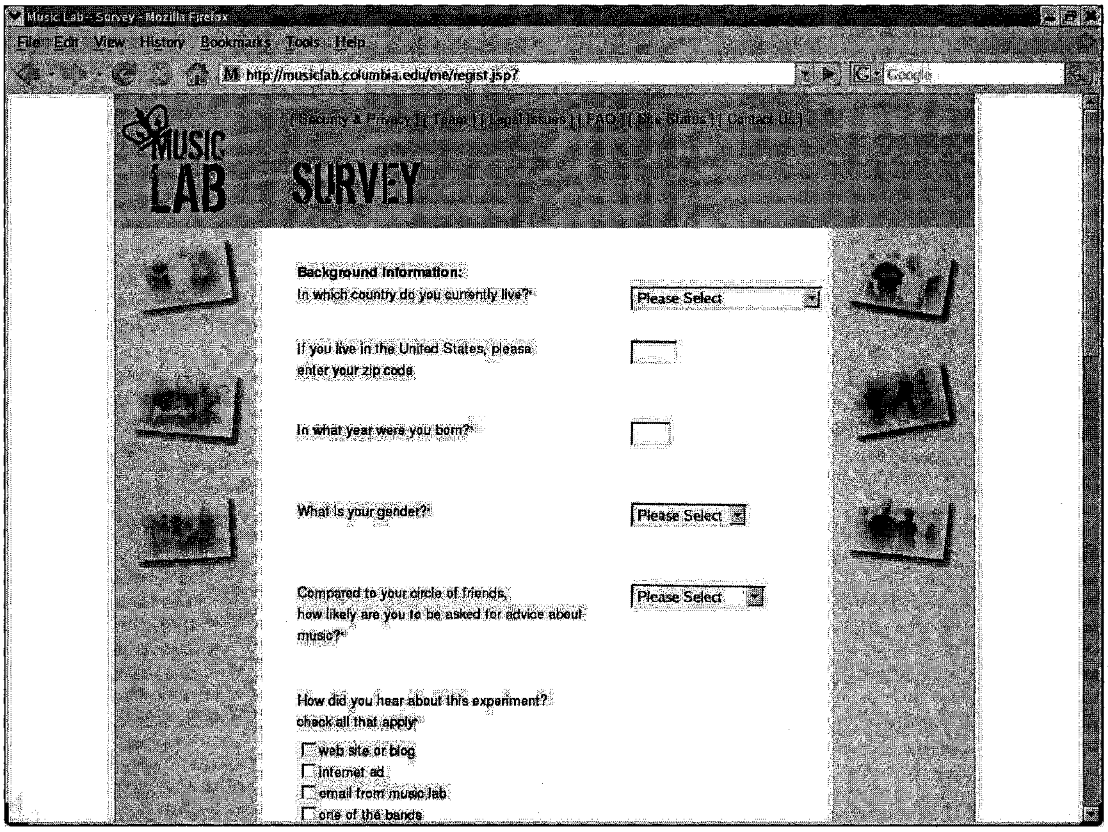


Figure 2.5: Survey screen from the website.

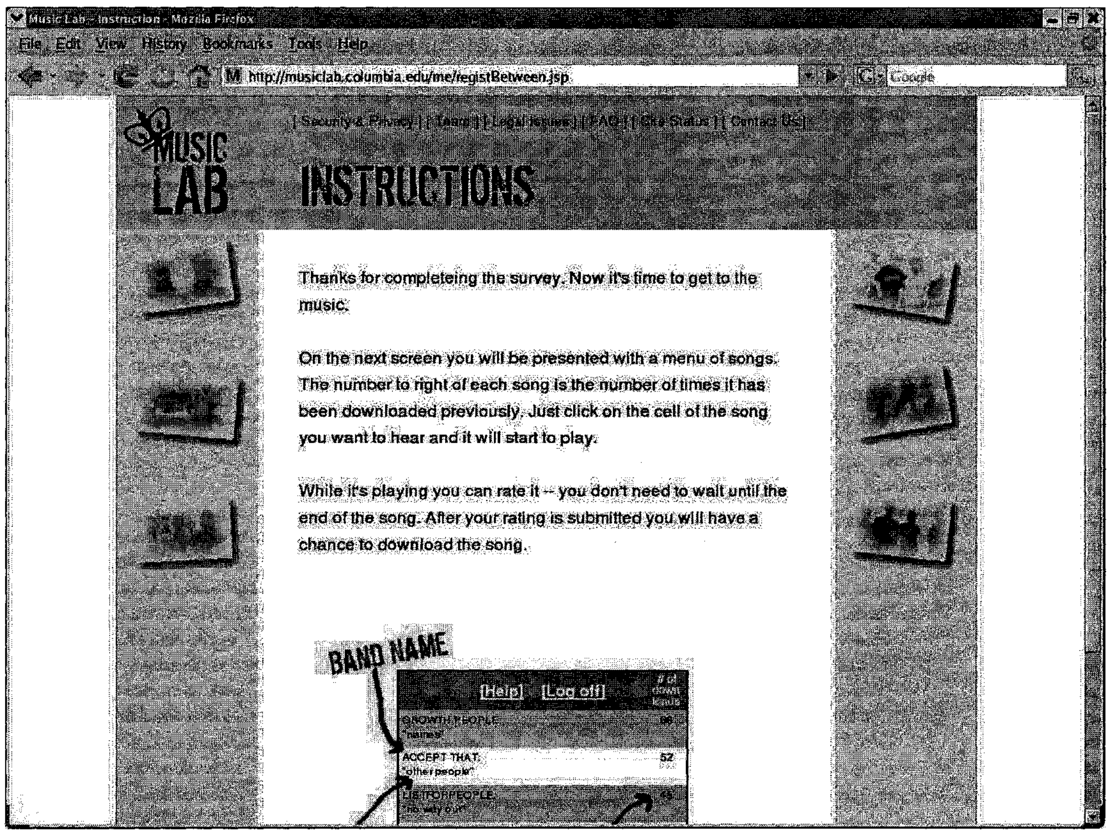


Figure 2.6: Instructions screen from the website.

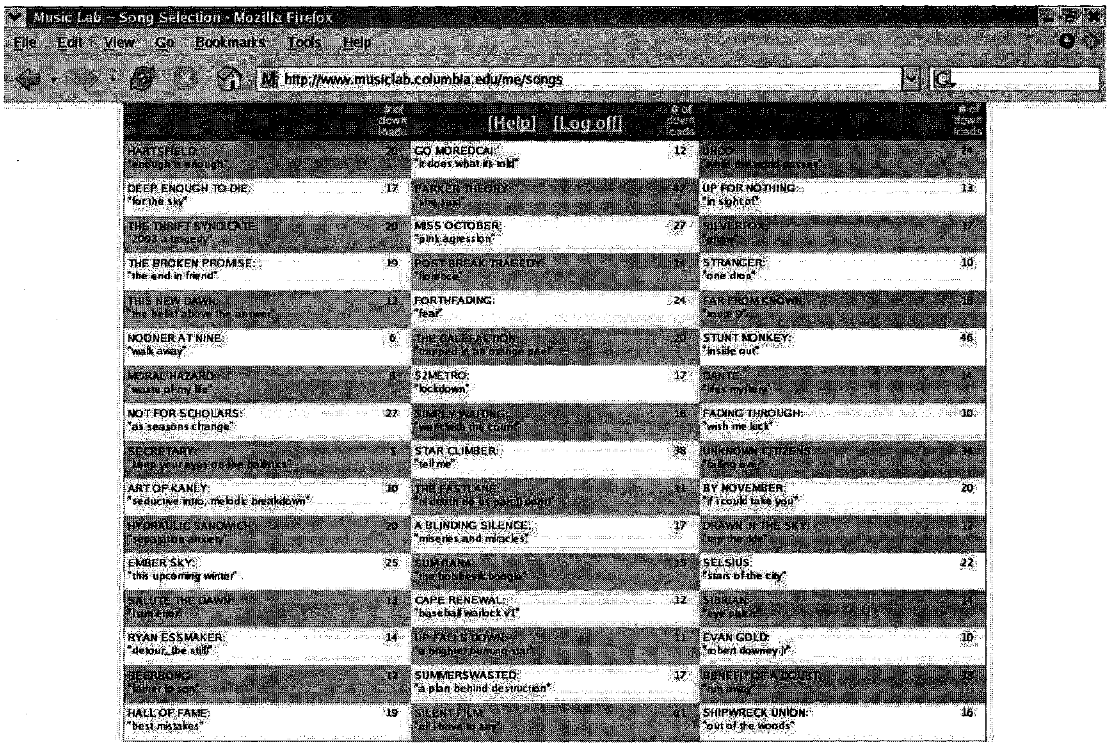


Figure 2.7: Screenshot of the song menu from a social influence world in experiment 1. The song menu in the independent condition (not shown) was identical except that the download counts to the right of each song were not present. In both conditions songs were presented to each participant in a random order.

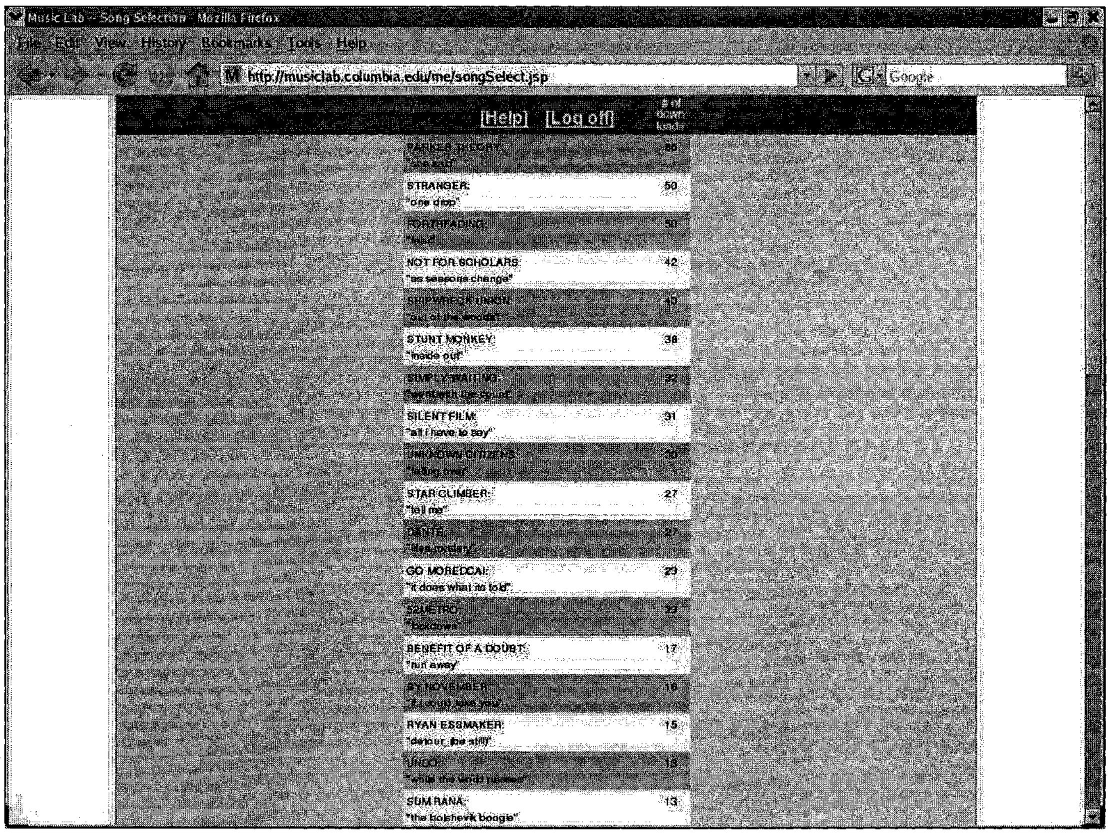


Figure 2.8: Screenshot of the song menu in a social influence world in experiments 2, 3, and 4. The song menu in the independent condition (not shown) was identical except that the download counts to the right of each song were not present. In the social influence worlds the songs were sorted by popularity and in the independent condition they were ordered randomly.

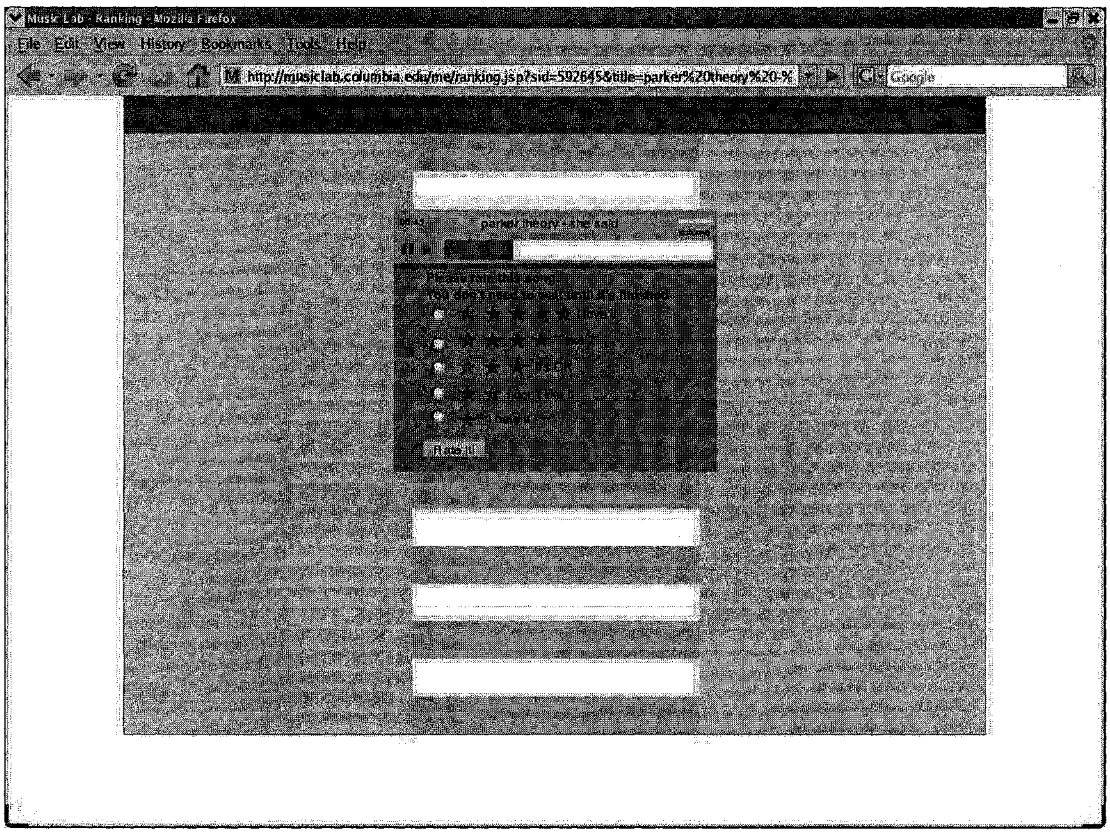


Figure 2.9: Screenshot of the listening screen. While a song was playing, subjects were required to rate it on a scale of 1 to 5 stars. This rating could be submitted before the song was finished playing.

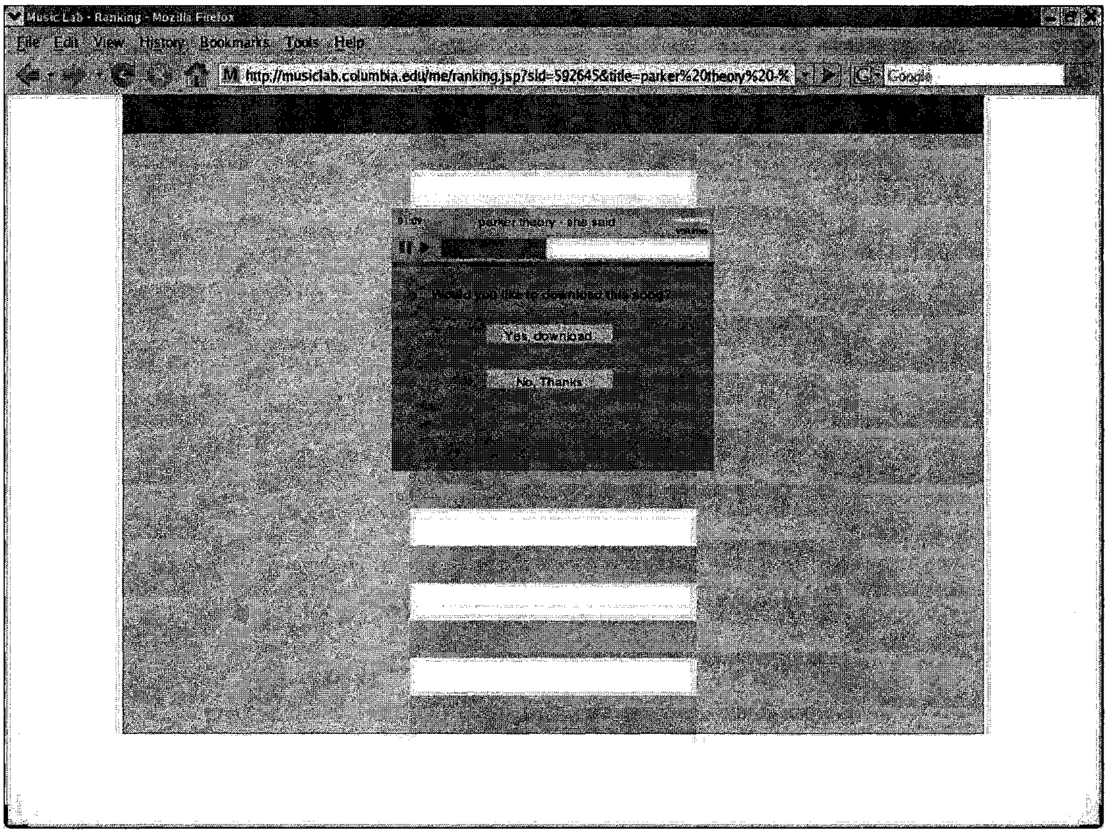


Figure 2.10: Screenshot of the download decision screen. After rating the song, subjects had to decide to download the song or not.

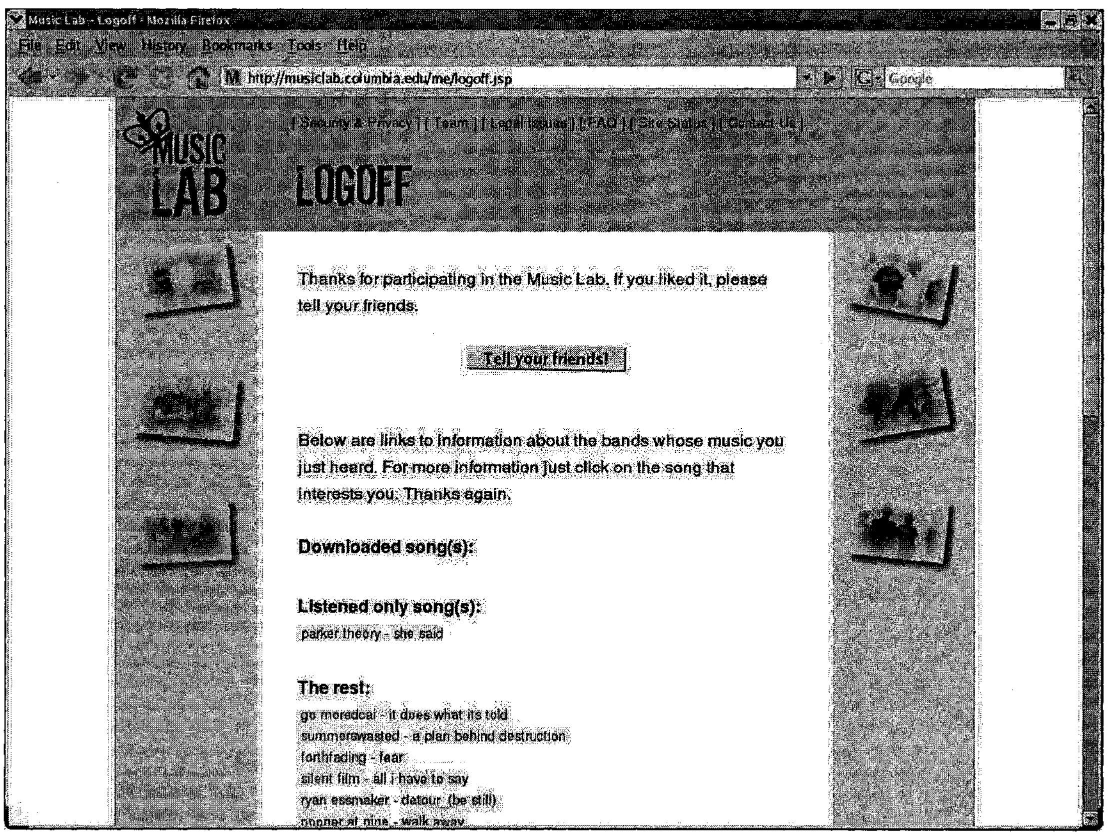


Figure 2.11: Logoff screen from the website.

table 2.1.

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 2.13. 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`.

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 2.14 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 2.1 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, Fig. 2.12). 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 (Dodds et al., 2003). Recruitment dynamics are presented in figure 2.15, 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.[5] The source of the spike around day 60 is unknown.[6] 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 2.1). 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.

[5]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.

[6]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.

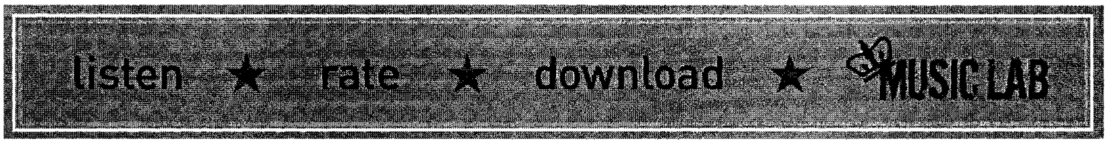
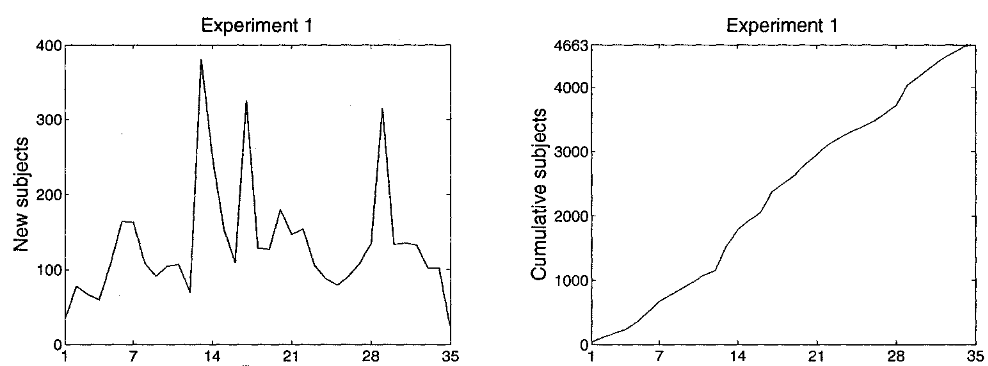


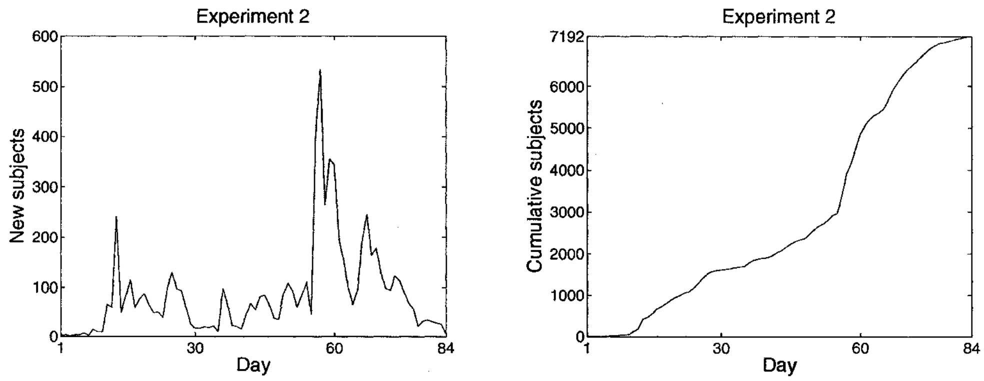
Figure 2.12: Banner advertisement used to recruit subjects from `http://www.bolt.com` for experiment 2.



(a) Subjects per day

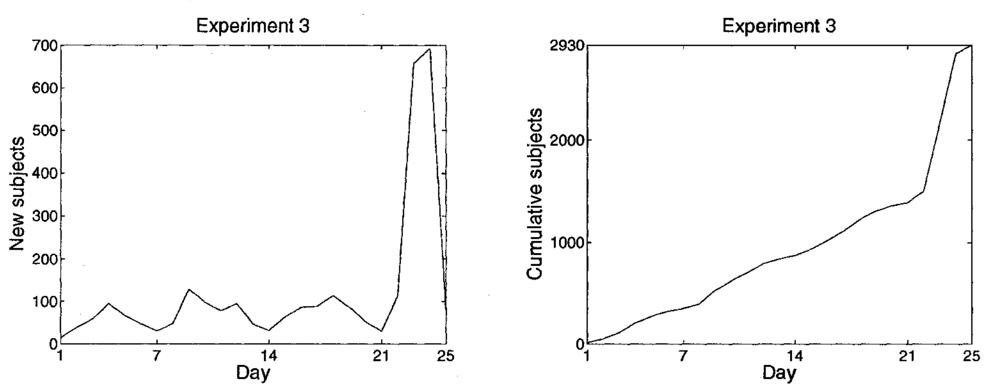
(b) Cumulative subjects

Figure 2.13: 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.



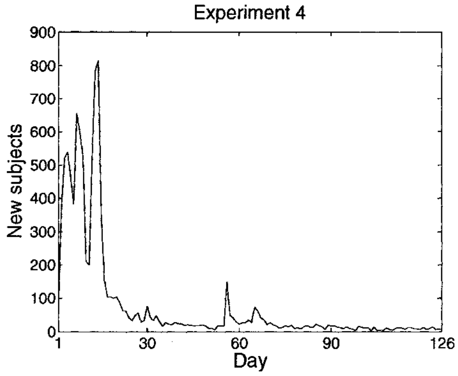
(a) Subjects per day (b) Cumulative subjects

Figure 2.14: 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`.

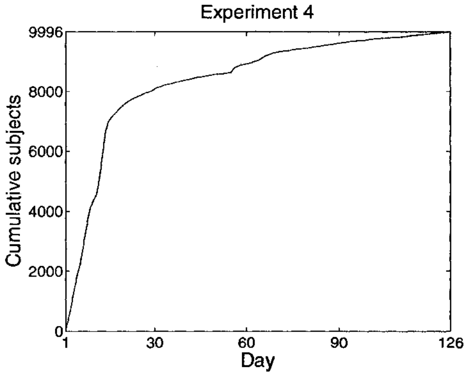


(a) Subjects per day (b) Cumulative subjects

Figure 2.15: 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 2.16: 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.

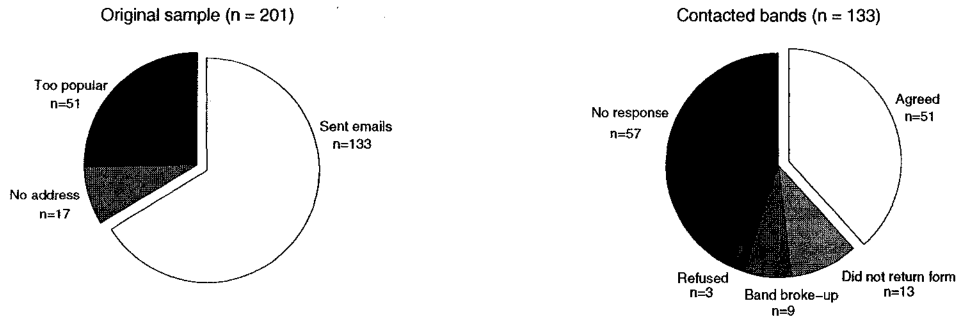
2.4 Finding the songs

The music for the experiment (see table 2.2) 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 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

Band name	Song name
52metro	Lockdown
A Blinding Silence	Miseries and Miracles
Art of Kanly	Seductive Intro, Melodic Breakdown
Beerbong	Father to Son
Benefit of a Doubt	Run Away
By November	If I Could Take You
Cape Renewal	Baseball Warlock v1
Dante	Life's Mystery
Deep Enough to Die	For the Sky
Drawn in the Sky	Tap the Ride
Ember Sky	This Upcoming Winter
Evan Gold	Robert Downey Jr.
Fading Through	Wish me Luck
Far from Known	Route 9
Forthfading	Fear
Go Mordecai	It Does What Its Told
Hall of Fame	Best Mistakes
Hartsfield	Enough is Enough
Hydraulic Sandwich	Separation Anxiety
Miss October	Pink Aggression
Moral Hazard	Waste of my Life
Nooner at Nine	Walk Away
Not for Scholars	As Seasons Change
Parker Theory	She Said
Post Break Tragedy	Florence
Ryan Essmaker	Detour_(Be Still)
Salute the Dawn	I am Error
Secretary	Keep Your Eyes on the Ballistics
Selsius	Stars of the City
Shipwreck Union	Out of the Woods
Sibrian	Eye Patch
Silent Film	All I have to Say
Silverfox	Gnaw
Simply Waiting	Went with the Count
Star Climber	Tell Me
Stranger	One Drop
Stunt Monkey	Inside Out
Sum Rana	The Bolshevik Boogie
Summerswasted	A Plan Behind Destruction
The Broken Promise	The End in Friend
The Calefaction	Trapped in an Orange Peel
The Fastlane	Til Death do us Part (I don't)
The Thrift Syndicate	2003 a Tragedy
This New Dawn	The Belief Above the Answer
Undo	While the World Passes
Unknown Citizens	Falling Over
Up Falls Down	A Brighter Burning Star
Up for Nothing	In Sight Of

Table 2.2: The 48 bands and songs used in the experiments.



(a) Original sample (n = 201) (b) Contacted bands (n = 133)

Figure 2.17: 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.

publicly available email address. The remaining 133 bands were contacted via email (results summarized in figure 2.17(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 2.17(b)). The email to the bands and band consent form are available in section A.2 and A.3.

Preliminary pilot testing revealed that, for the song menu used in experiment 1 (figure 2.7), 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 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

		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 2.3: 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.

Remnant Solder), an imaginary band (Peter on Fire), and an extremely well known band (U2).[7] Table 2.3 and figure 2.18 show that some subjects reported being familiar with the three potential bands, but these recognition rates were no higher than for the imaginary band.[8] 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.

[7]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 (Goldstein and Gigerenzer, 2002).

[8]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.

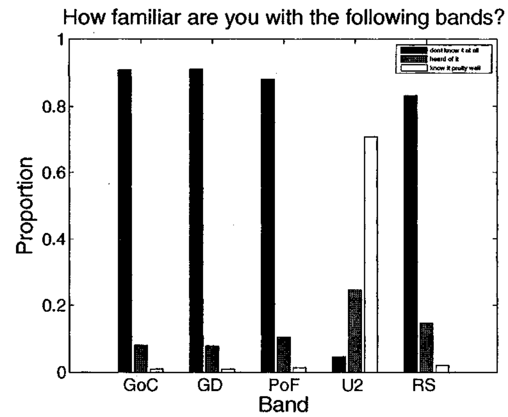


Figure 2.18: Comparing the popularity of the potential bands from our sample to a fake band. Participants reported being about as familiar with an fake band (Peter on Fire [PoF]) as three potential bands from our sample: Guys on Couch [GoC], Grover Dill [GD], and Remnant Soldier [RS]. The higher recognition rate for Remnant Soldier is likely a question ordering effect—it was asked immediately after the well known band U2.

2.5 Data quality protocols

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 (Nosek et al., 2002). 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.

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 (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.

Now that the appropriate background has been provided about the experimental set-up, we turn our attention to the results of the four experiments.

Chapter 3

Experiments 1 and 2

3.1 Introduction

The conceptual model developed in chapter 1, and shown in figure 3.1, makes a number of clear, falsifiable predictions. First, it suggests that outcomes in the presence of social influence will be more unequal and unpredictable than outcomes where individuals act independently. Further, the model suggests that increasing the strength of the social influence at the individual level should increase the amount of inequality and unpredictability at the collective level. We tested these predictions in a series of two experiments which differed only in the amount of social influence that we attempted to impose on the participants. In experiment 1, the song menu was arranged in a 16×3 unsorted grid (figure 3.2) while in experiment 2 the song menu was arranged in one-column list sorted by popularity (figure 3.3).[1]

[1]Some of the results in this chapter have been published in a previous paper co-authored with Peter Dodds and Duncan Watts (Salganik et al., 2006).

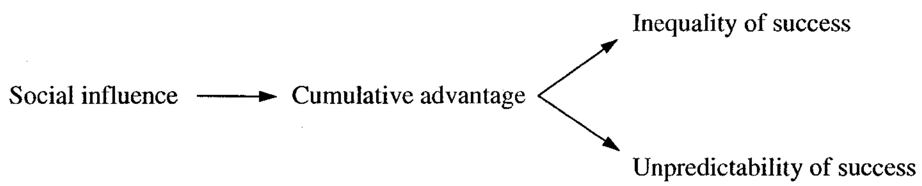


Figure 3.1: Schematic of the conceptual model. The inequality and unpredictability of success, which seem to be contradictory at the aggregate level, can both arise from social influence at the level of the individual leading to cumulative advantage in the success of objects.

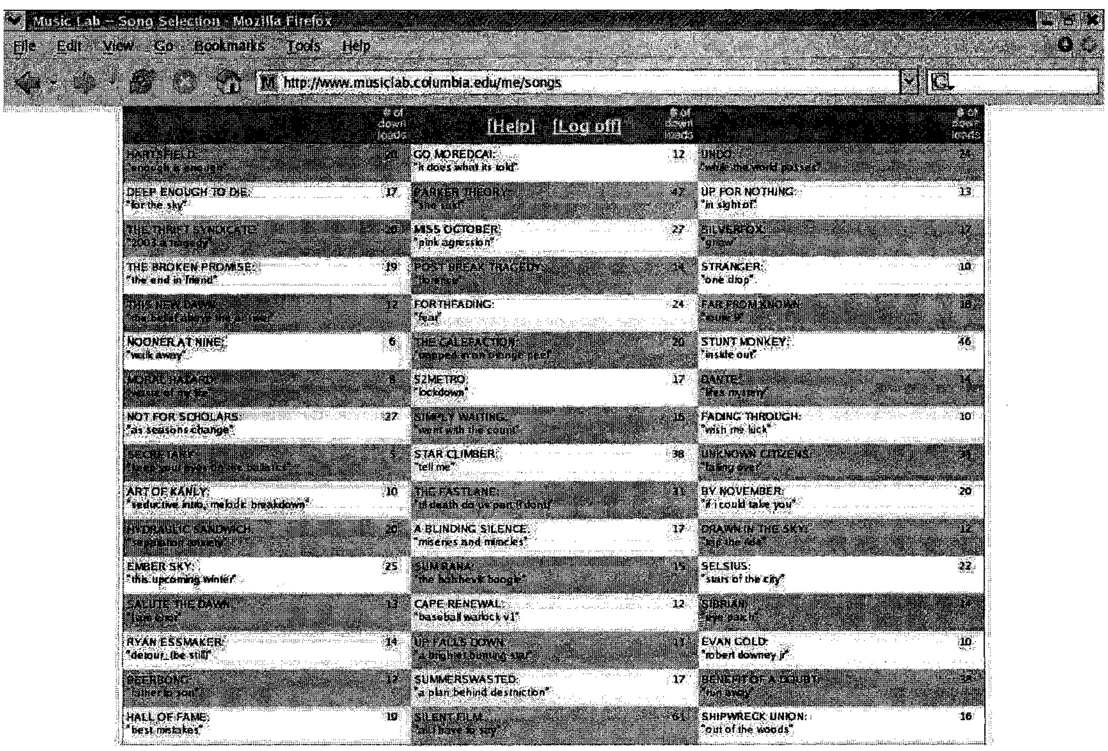


Figure 3.2: Screenshot of the song menu from a social influence world in experiment 1. The song menu for participants in the independent condition (not shown) was identical except that the download counts to the right of each song were not present. In both conditions songs were presented to each participant in a random order.

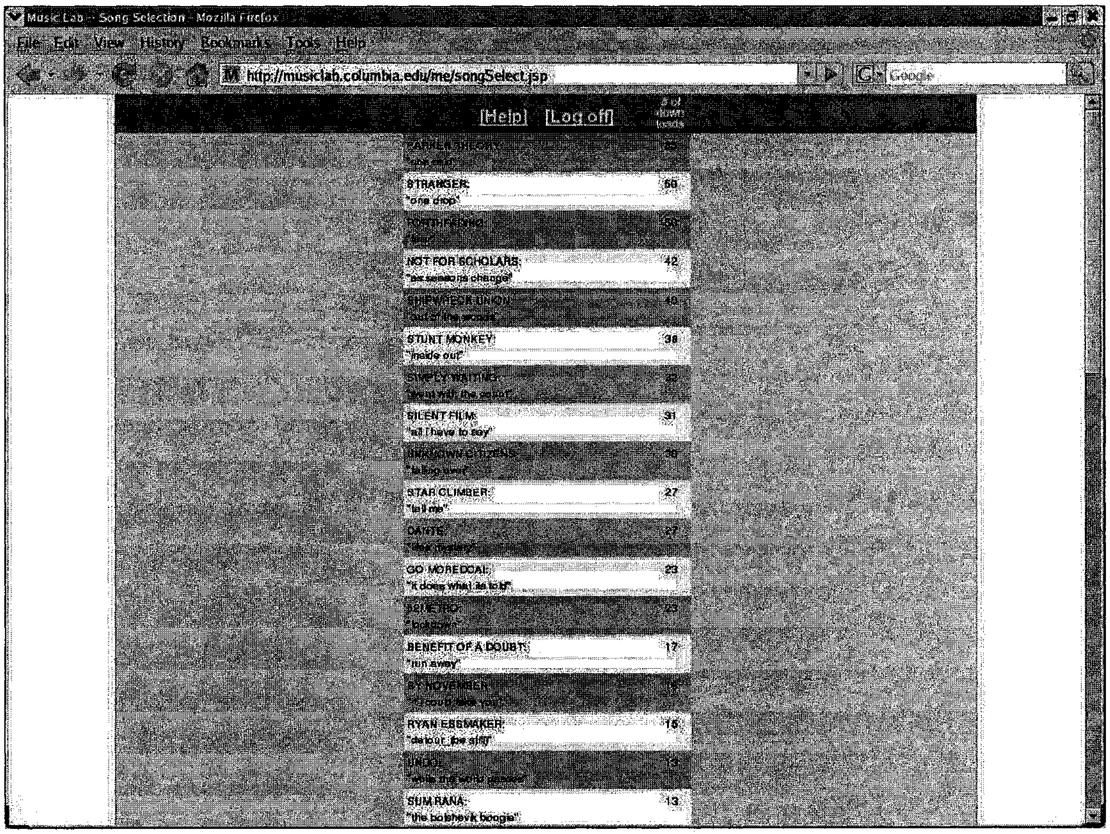


Figure 3.3: Screenshot of the song menu in the social influence worlds in experiments 2. The song menu for participants in the independent condition (not shown) was identical except that the download counts to the right of each song were removed. In the social influence worlds the songs were sorted by popularity and in the independent condition they were ordered randomly.

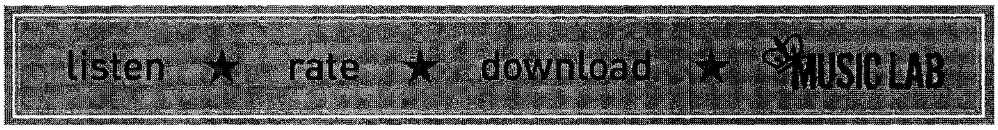


Figure 3.4: Banner advertisement used to recruit subjects from http://www.bolt.com for experiment 2.

3.1.1 Subject recruitment

Experiment 1 took place from October 7, 2004 to December 15, 2004 (69 days) and involved 7,149 subjects. 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. Most subjects for both experiments were recruited from http://www.bolt.com, a website popular with teens and young adults from the United States. The recruitment involved links on the "music" and "free-stuff" sections of the website, a "BoltNote" that was sent to all members, and banner ads run in all section of site (figure 3.4). Demographics about these subjects are presented in table 3.1. We note that there was a change in percentage of females from experiment 1 to experiment 2. Subjects in both experiments were drawn from http://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 the "BoltNote" and the banner ads. Another potential reason for the difference is that while experiment 1 was underway, the project was mentioned on the popular blog http://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.

Summary statistics about participant behavior is presented in table 3.2. Figure 3.5 plots the distribution of listens per subject in experiment 1. A surprisingly high number of participants did not listen to any songs. The exact reason for this is

Category	Experiment 1 ($n = 7,149$) (% of participants)	Experiment 2 ($n = 7,192$) (% of participants)
Female	36.4	73.9
Broadband connection	74.1	69.0
Has downloaded music from other sites	60.4	62.4
Country of Residence		
UNITED STATES	79.8	81.8
CANADA	4.5	4.4
UNITED KINGDOM	4.4	4.7
OTHER	11.0	9.1
Age		
14 AND YOUNGER	11.5	16.0
15 TO 17	27.8	34.9
18 TO 24	38.5	39.2
25 AND OLDER	22.3	9.9

Table 3.1: Descriptive statistics about the participants in experiments 1 and 2 most of whom were recruited from `www.bolt.com`.

unknown, but we suspect that it was because subjects did not recognize any of the songs.[2] These subjects who did not listen to any songs did not effect our results, but they did reduce our effective number of participants. Figure 3.6 plots the distribution of downloads per subject in experiment 1. Figures 3.7 and 3.8 plot the same information for experiment 2. Again, a large number of participants did not listen to any songs.

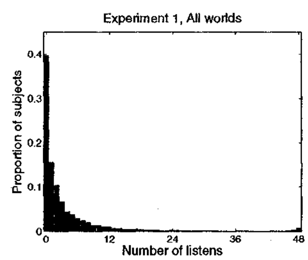
3.2 Experiment 1

During experiment 1, 7,149 subjects were assigned to either one of eight social influence worlds or the independent condition. This design yielded ~ 700 subjects in each influence world and $\sim 1,400$ subjects in the independent condition (the reason for having twice as many subjects in the independent condition will become

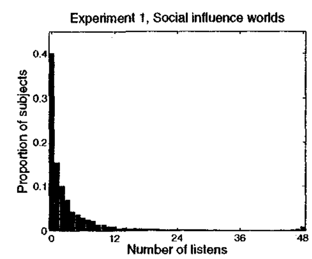
[2]We emailed some of these non-listening participants to ask them why, but they did not respond to our email.

	Experiment 1			Experiment 2		
	Influence ($n = 5,708$)	Independent ($n = 1,441$)	Total ($n = 7,149$)	Influence ($n = 5,746$)	Independent ($n = 1,446$)	Total ($n = 7,192$)
Number of listens	21,971	5,394	27,365	20,217	5,643	25,860
Mean per subject	3.8	3.7	3.8	3.5	3.9	3.6
Median per subject	1	1	1	1	1	1
Number of downloads	6,626	1,578	8,203	8,106	2,192	10,298
Mean per subject	1.2	1.1	1.1	1.4	1.5	1.4
Median per subject	0	0	0	0	0	0
Pr[download | listen]	0.302	0.293	0.300	0.401	0.388	0.398
Average rating (# of stars)	3.0	2.9	3.0	3.2	3.2	3.2

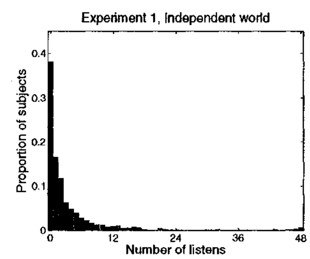
Table 3.2: Descriptive statistics of subject behavior in experiments 1 and 2.



(a) Exp 1, all worlds

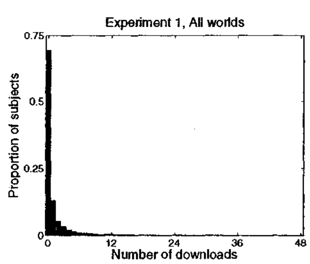


(b) Exp 1, social influence worlds

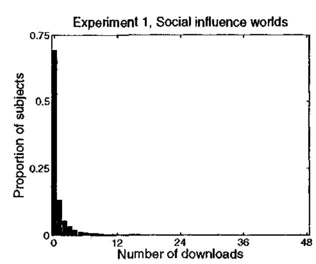


(c) Exp 1, independent world

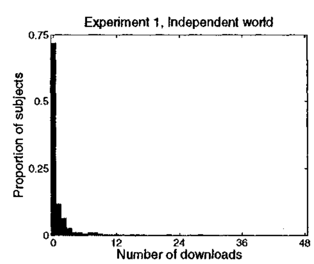
Figure 3.5: Distribution of the number of listens per participant in experiment 1. Results are also presented for just participants in the social influence worlds and just participants in the independent world.



(a) Exp 1, all worlds

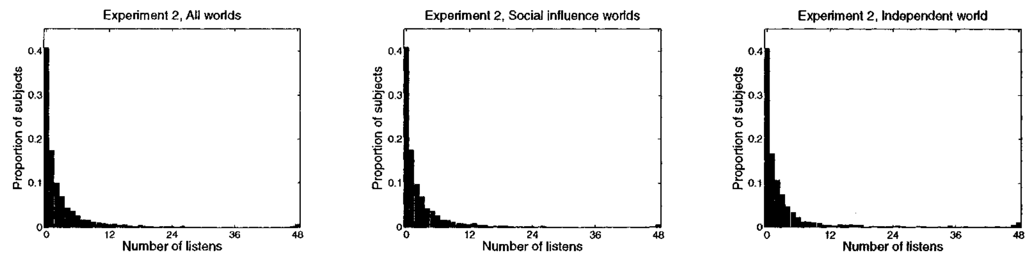


(b) Exp 1, social influence worlds



(c) Exp 1, independent world

Figure 3.6: Distribution of the number of downloads per participant in experiment 1. Results are also presented for just participants in the social influence worlds and just participants in the independent world.

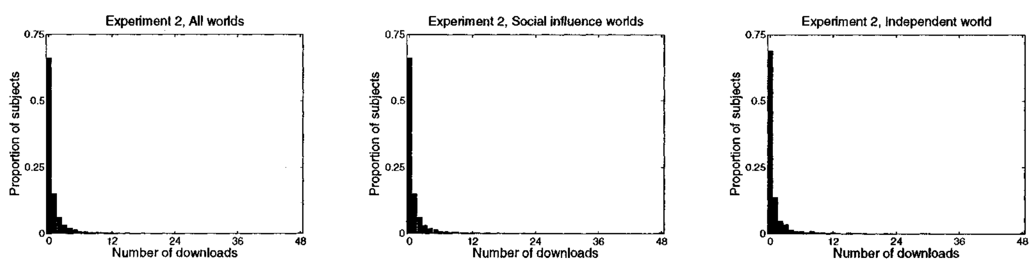


(a) Exp 2, all worlds

(b) Exp 2, social influence worlds

(c) Exp 2, independent world

Figure 3.7: Distribution of the number of listens per participant in experiment 2. Results are also presented for just participants in the social influence worlds and just participants in the independent world.



(a) Exp 2, all worlds

(b) Exp 2, social influence worlds

(c) Exp 2, independent world

Figure 3.8: Distribution of the number of downloads per participant in experiment 2. Results are also presented for just participants in the social influence worlds and just participants in the independent world.

clear when we discuss our measure of unpredictability). All songs started with zero downloads and the download counts were updated in real-time as the experiment progressed.

3.2.1 Individual level behavior

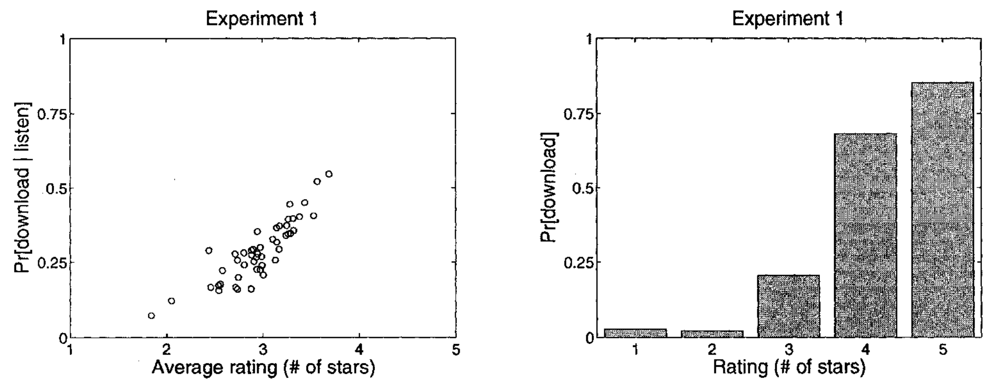
Within our experimental framework, there are two main ways that social influence can affect a subject's behavior because there are two main decisions that the subject must make before a song is downloaded. First, the subject must decide which of the 48 available songs to listen to. Then, while listening to a song, the subject must decide whether or not to download the song. Information about the popularity of the songs could affect both of these decisions. We will focus our attention on the effect of popularity on subjects' listening decisions.[3]

We measure the success of a song in a given world in terms of its market share, m_i, defined to be the fraction of all downloads that belong to that song,

$$m_i = \frac{d_i}{\sum_{k=1}^{S} d_k} \tag{3.1}$$

where d_i is the number of downloads for song i and S is the number of songs. This definition of success is based on the subjects' behavior rather than their self-reported liking of the songs (ranging from 1 to 5 stars). It turns out, however, that these two measures are largely consistent. Figure 3.9 shows that, in experiment 1, songs which received higher average ratings (measured in stars) had higher probabilities

[3]It is more difficult to detect social influence on the download decision because subjects chose which songs they want to listen to and thus self-selected into the download decision. For example, imagine that we observed that when a song was more popular it had a higher probability of converting a listen to a download than when that song was less popular. This would seem to suggest that there was social influence on the download decision. However, this result could be because the subjects who chose to listen to the most popular songs were different from those who chose to listen to the least popular songs. Thus, what appears to be an effect of the popularity of the song could have been a compositional effect. This compositional effect does not confound our experiment, just our attempt to measure social influence on download decisions. This issue will be further address in chapter 6.



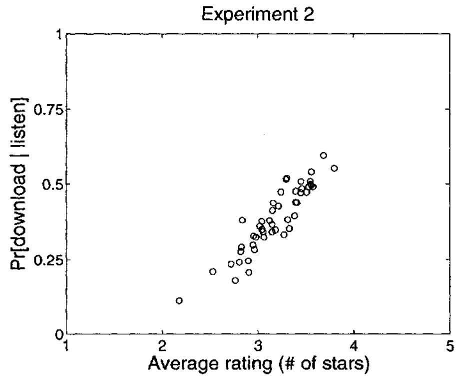
(a) Comparing behavioral and self-report data (b) Comparing behavioral and self-report data

Figure 3.9: Plots comparing the rating decisions and download decisions in experiment 1. These results suggest that the behavioral and self-reported measures are consistent.

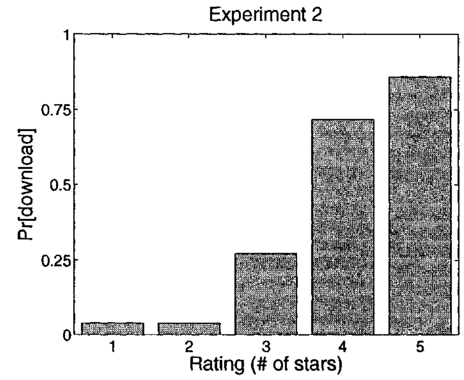
that a listen would result in a download ($r = 0.87$). Further, figure 3.9(b) shows that, in experiment 1, the higher rating a subject gave a song, the more likely that the subject downloaded the song. Results from experiment 2 are essentially identical (figure 3.10). Overall, the similarity between these two measures gives us confidence that our behavioral measure is meaningful and thus a reasonable measure of success.

Given this measure of success, at the time each subject participated, every song in their world had a specific market share and market rank (for example, the song with the highest market share has a market rank of 1). We can measure the influence of market rank on subjects' listening decisions by calculating the probability that a subject in the social influence condition chose to listen to the song of a given market rank (independent of which song occupied that rank at the time). For example, figure 3.11 shows that in the influence worlds each subject had a 20% chance of listening to whatever was the most popular song but only a 7% chance of listening to whatever was the least popular song.[4] Such differences did not occur in the inde-

[4]The values presented in figure 3.11 have been smoothed to aid visualization. A description of the the smoothing procedure and plots of unsmoothed data are presented in appendix B.3. Also, it



(a) Comparing behavioral and self-report data



(b) Comparing behavioral and self-report data

Figure 3.10: Plots comparing the rating decisions and download decisions in experiment 2. These results suggest that the behavioral and self-reported measures are consistent.

pendent condition. Also, figure 3.11 shows a clear "top 10 effect"; that is, subjects in the social influence condition were not affected by market rank except for the top 10 spots. Within this top 10, there is a clear pattern of subjects choosing to listen to higher ranked songs. It is interesting to note that this top 10 effect occurred even though the song menu (see figure 3.2) was not sorted by popularity. Given that social influence at the individual level occurred, we now explore its effect on the aggregate level features of the market for songs.

3.2.2 Inequality

Regarding the inequality of success, the conceptual model outlined above suggests that the success of the songs in the social influence worlds should exhibit more inequality than in the independent condition. Equality of success is easy to define, but there are many different ways to quantify the amount of inequality (Allison, 1978;

is important to note that the probabilities in this plot do not sum to 1 because the figure has 95 data points: the 48 possible ranks plus the 47 possible ties (e.g. two songs that are tied in first place will both be assigned a rank of 1.5).

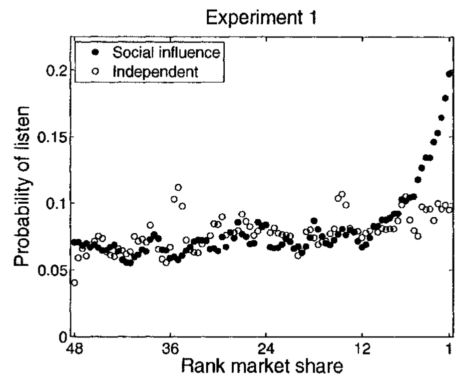


Figure 3.11: Probability that a subject in each condition listened to a song of a given market rank in experiment 1. Note that this figure is smoothed for ease of presentation. Unsmoothed results, which are qualitatively the same, are presented in figure B.22.

Coulter, 1989). We present our results using one of the most common metrics, the Gini coefficient, G, defined as follows,

$$G = \frac{\frac{1}{S^2}\sum_{i=1}^{S}\sum_{j=1}^{S}|m_i - m_j|}{2 \cdot \frac{\sum_{i=k}^{S} m_k}{S}} \tag{3.2}$$

where m_i is the market share of song i and S is the number of songs. The Gini coefficient can be interpreted as the expected difference in market share between two randomly chosen songs scaled so that it falls between 0 (complete equality) and 1 (maximal inequality).[5] In figure 3.12 we see that the Gini coefficients for the eight social influence worlds ranged from 0.28 to 0.42 with a median value of 0.33, while the independent condition had a Gini coefficient of 0.26.

In order to make a fair comparison between the eight influence worlds which had ~ 700 subjects and the independent condition which had $\sim 1,400$ subjects, we

[5]When choosing our measure of inequality, we considered other measures that are scale invariant and obey the principle of transfers (Coulter, 1989). Thus, in addition to the Gini coefficient, we checked our hypothesis using the coefficient of variation and the Herfindahl index; the results were qualitatively unchanged no matter which metric we used (see appendix B.2). We also considered the market concentration in the top 5 songs, a measure which does not obey the principle of transfers, but is used frequently. Again the results are qualitatively unchanged (see appendix B.2). More detailed descriptions of these measures, and others, can be found in Allison (1978) and Coulter (1989).

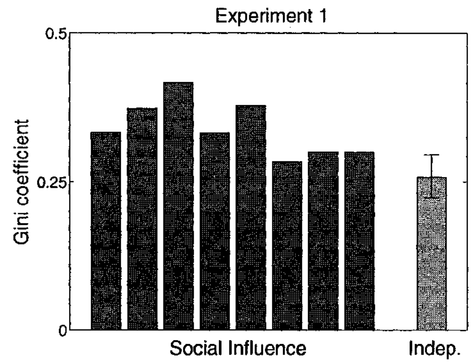


Figure 3.12: Inequality of success in experiment 1.

randomly subsampled the independent condition to produce an independent world of ~ 700 subjects. We then repeated this procedure 1,000 times and measured the Gini coefficient in each subsample. The value presented in figure 3.12 is the average of these 1,000 Gini coefficients and the whisker represents the interval in which 95% of the replicates fall.[6]

In order to assess the difference in inequality across conditions, we first note that the independent world is more equal than all of the social influence worlds. To get a sense of the size of the differences across conditions, we can make a comparison to the income inequality in different countries. The difference between the conditions is similar to the difference between income inequality in Western European countries like Spain, the Netherlands, and France ($G \approx 0.33$) and Scandinavian countries like Sweden, Denmark, and Norway ($G \approx 0.24$) (UNDP, 2004). Finally, we can preform a test of statistical significance using the distribution of values that was created by the subsampling procedure. The difference between a randomly chosen Gini coefficient from one of the eight influence worlds and a randomly chosen replicate Gini coefficient from the independent world was less than 0 with $p < 0.01$ in experiment 1. Thus,

[6]This subsampling procedure does not change any of our substantive conclusions because the Gini coefficient of the entire independent condition ($n \approx 1,400$) was 0.24, which is lower than our presented value.

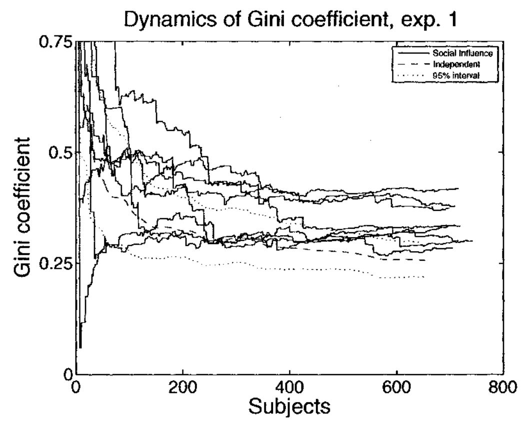


Figure 3.13: Dynamics of the Gini coefficient in experiment 1.

the difference in observed Gini coefficients between the two conditions is statistical significant.

Finally, we can examine the dynamics of the Gini coefficient as the experiment progresses (figure 3.13 and figure 3.14).[7] The final values of each trajectory are the values reported in 3.12. The Gini coefficients were relatively stable indicating that we probably would not have observed substantially different results with more subjects.

3.2.3 Unpredictability

Having shown that social influence increased the inequality of success, we now turn to the second hypothesis—that social influence causes the success of songs to be more unpredictable. Here we conceptualize unpredictability in terms of the variation in success of a given song across different realizations of the same process. If the outcome is the same in all realizations then it is, at least in theory, predictable, and its unpredictability is defined to be 0. This does not mean that in practice it will

[7] A careful review of figure 3.14 reveals that the trajectory of the Gini coefficient in world 1 is unusual in that it drops very quickly and then increases over time, a pattern not seen in the other seven worlds. The reason for these dynamics is that the 7th user in this world (*uid* = 103) downloaded almost every song. This user action lead to a dramatic decrease in inequality which took many downloads to wash out.

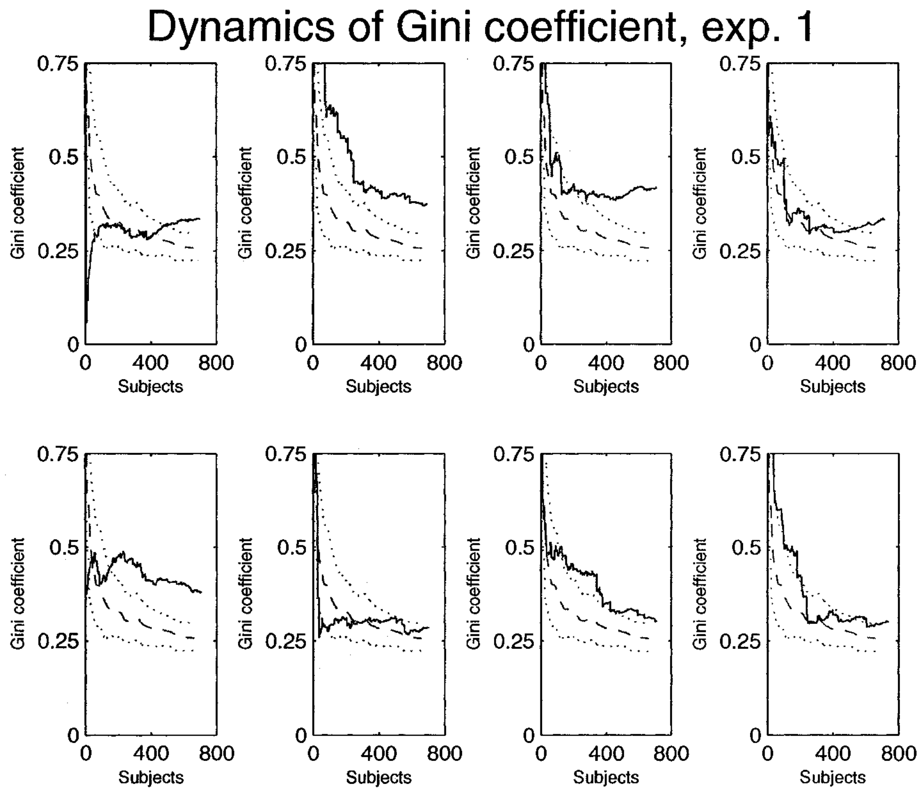


Figure 3.14: Dynamics of the Gini coefficient in experiment 1 for each world. Solid lines are the social influence worlds; the dashed line is the independent world; and the dotted lines represent 95% Monte Carlo intervals around the average independent world.

be easy to predict the success of the song, but since its success is the same in each realization, precise prediction is possible, at least in principle. However, if outcomes differ between realizations, then precise prediction becomes impossible and the larger the differences between realizations, the more unpredictable a song's success.[8]

In an attempt to capture the magnitude of these realization-to-realization differences in success, we define the unpredictability of a specific song, u_i, to be the average difference in market share between two randomly chosen success outcomes. That is,

$$u_i = \frac{\sum_{j=1}^{R} \sum_{k=j+1}^{R} \mid m_{i,j} - m_{i,k} \mid}{\binom{R}{2}} \tag{3.3}$$

where $m_{i,j}$ is the market share of song i in realization j, and $\binom{R}{2}$ is the number of pairs of realizations. A larger value of u_i implies more difference between two randomly chosen realizations, and thus greater unpredictability of outcomes. For example, we can compare the unpredictability of success in the social influence condition for two songs "Inside Out" by Stunt Monkey and "Fear" by Forthfading. Figure 3.15 plots the eight different social influence outcomes for each song. Visual inspection reveals, first, that the outcome of each song exhibited some inherent unpredictability. Visual inspection also suggests that because of the larger range of outcomes for "Fear," the average difference between two randomly chosen outcomes would be larger for this song. This intuition is captured by our unpredictability measure u_i (defined in equation 3.3) which yields values of $u = 0.014$ for "Inside Out" and a larger value of $u = 0.021$ for "Fear."

Once we have measured the unpredictability of a song, we must calculate the unpredictability for an entire experimental condition. Thus, we define U to be the

[8]Returning again to the example of a die. If one has a weighted die which always produces the same value, it is possible, although not necessarily easy, to determine what value the die will produce before rolling it. If one has a fair die, however, no amount of information about the physical characteristics of the die enables precise predictions because the outcomes vary from roll to roll.

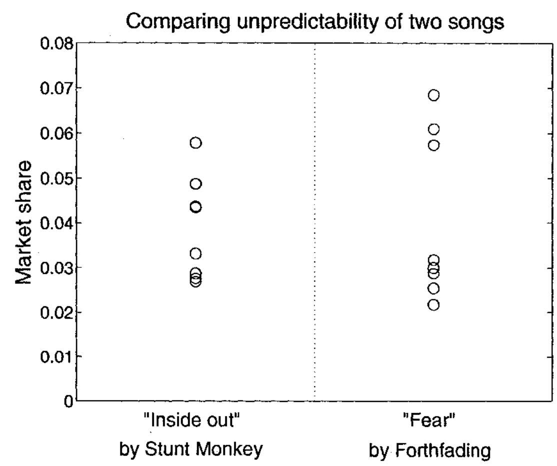


Figure 3.15: Comparing the unpredictability in the social influence condition of two songs in experiment 1. Final market shares from all eight social influence worlds are plotted for "Inside Out" by Stunt Money and "Fear" by Forthfading. The eight different worlds produced different success outcomes for each song suggesting that success was inherently unpredictable. Further, our song-level measure of unpredictability, u_i, captures the fact that the outcomes for "Fear" were more unpredictable than for "Inside Out".

average unpredictability of all songs in a given condition,

$$U = \frac{\sum_{i=1}^{S} u_i}{S} \tag{3.4}$$

where S is the number of songs.

One difficultly with this measure is that in the independent condition we have only one world. However, as noted previously, it had twice as many subjects as each social influence world. Thus, for the independent condition, we randomly split the subjects into two independent realizations and calculated u_i and U with these two realizations.[9] We repeated this splitting procedure 1,000 times and produced a distribution of replicate values of U in the independent condition. In the social influence condition, calculated U values for the 28 $(\frac{8\times7}{2})$ possible pairs of influence worlds. Figure 3.16 shows that, as predicted, the unpredictability, U, is greater in

[9]Since the behavior of the subjects in the independent condition was not affected by the behavior of previous participants, it is irrelevant how we split the population.

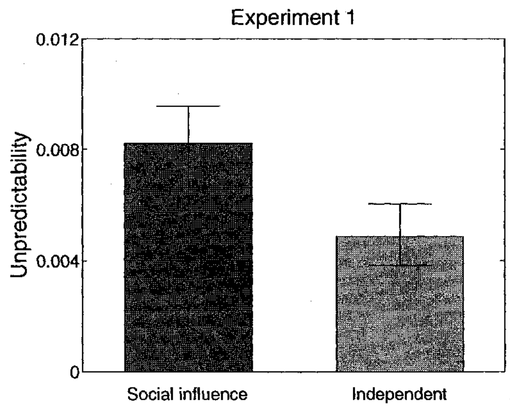


Figure 3.16: Unpredictability in experiment 1.

the social influence condition than the independent condition.[10] The unpredictability described here is caused by social influence and therefore is inherent to the aggregation process; in the same way that rolling a die is stochastic and unpredictable, so too is the success of a given song.

To calculate a measure of statistical significance we compared the distribution of replicate values. The difference between the unpredictability based on a randomly chosen pair of social influence worlds and the unpredictability based on a random split of the independent world was less than 0 with a probability of $p < 0.01$.

Finally, we can examine the dynamics of the unpredictability U as the experiment progresses (figure 3.17). The final values of each trajectory are the values reported in figure 3.16. As with our measure of inequality, the unpredictability was relatively stable indicating that we probably would not have observed substantially different results with more subjects.

[10]The reason that the unpredictability of the independent world is not 0 is because we are comparing random splits of a group of about 1,400 people. The law of large numbers suggests that as the number of subjects in the independent condition gets large, the difference between success outcomes of two random splits will approach 0. This same argument does not apply to the social information worlds because those subjects are not acting independently and so the law of large numbers does not apply.

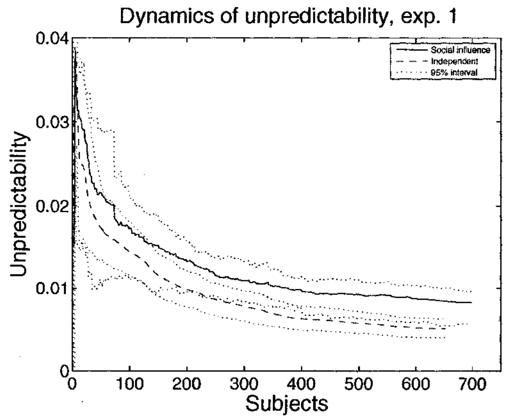


Figure 3.17: Dynamics of unpredictability in experiment 1.

3.2.4 Conclusion

Overall, the results from experiment 1 support the argument that social influence at the level of the individual causes both inequality and unpredictability of success. However, the differences between outcomes in the social influence and independent conditions, while real, are relatively small. One might doubt, therefore, that the process of social influence leading to cumulative advantage could actually be responsible for the large inequality and unpredictability that are observed in real cultural markets. However, our experiment clearly differs from real cultural markets in a number of ways. For example, the social influence in experiment 1 was quite weak because the signal of popularity was not very salient—a small number to the right of each song—and was easily ignored by subjects who were making anonymous decisions. That we were able to observe any effects at all, even under these extremely mild conditions, suggests the effects under stronger conditions could be substantial. Again, we note that our objective for experiment 1 was not to exactly replicate the conditions of the real world, but simply to measure the impact of social influence under controlled conditions (Zelditch, 1969; Lucas, 2003). Nevertheless, an interesting question arises: to what extent would increasing the social influence at the individual level affect the inequality and unpredictability in our controlled music market?

3.3 Experiment 2: Increasing the social influence

To explore the effects of increased social influence on aggregate outcomes, we re-designed the song menu used in experiment 1 in an attempt to increase the amount of social influence. Recall that in experiment 1, subjects were presented songs in three columns in a random order (figure 3.2). In experiment 2, subjects in the social influence condition were presented the songs in one column sorted by popularity (figure 3.3) while subjects in the independent condition were also presented the songs in one column, but in a random order (not sorted by popularity) and without the download counts. If several songs shared the same number of downloads, the ordering of the songs was determined randomly for each subject. This new song menu design was intended to make the popularity of the songs more salient and hence increase the amount of social influence on individual decisions.

Experiment 2 was conducted from December 15, 2004 until March 8, 2005 (83 days) during which time a total of 7,192 subjects were assigned to either one of eight social influence worlds or the one independent condition. As in experiment 1, this designed yielded ~ 700 subjects in each influence world and $\sim 1,400$ subjects in the independent condition. Also, as with experiment 1, most subjects were recruited from the website `www.bolt.com` and demographics are presented in table 3.1.[11] Therefore,

[11]As described previously, one thing to note is that there was a change in percentage of females from experiment 1 to experiment 2 (36.4 to 73.9). Subjects in both experiments were drawn from `http://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. Another potential reason for the difference is that while experiment 1 was underway, the project was mentioned on the popular blog `http://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.

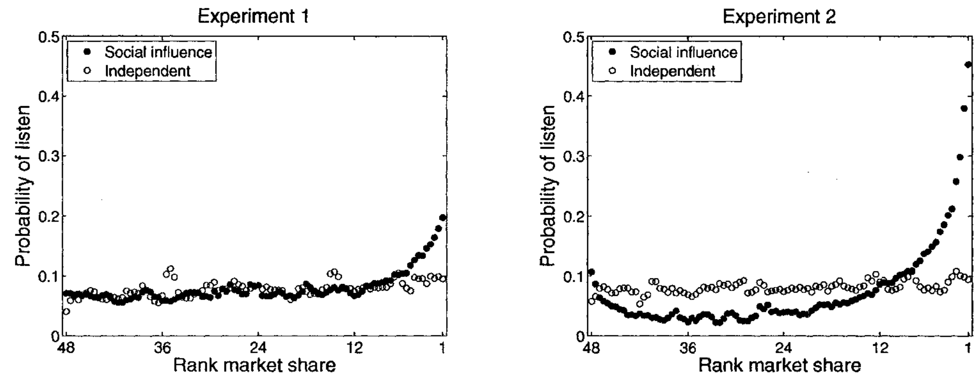
since subjects were drawn from the same source and they were evaluating the same set of songs, any difference between experiment 1 and experiment 2 can likely be attributed to the change in the layout of the song menu.[12] As in experiment 1, the download counts for all songs in all worlds started at zero and were updated in real-time.

3.3.1 Individual level behavior

First, we can test whether our intervention, changing the song menu from a unordered grid to a ranked list, had the effect of increasing social influence on subjects' decisions. We tested for this effect, as we did in the analysis of experiment 1, by calculating the probability that a subject in the social influence condition listened to a song of a given market rank. In figure 3.18 we can see that in both experiments there was a clear pattern of subjects choosing to listen to whichever songs had higher market rank and that this effect was stronger in experiment 2. For example, in experiment 1 subjects were about 3 times more likely to listen to the most popular song than a mid-level song; in experiment 2, subjects were about 8 times more likely.[13] From these results we can conclude that changing the song menu layout had the effect of increasing social influence at the individual level.

[12]In the future, researchers seeking to explore how alternate layouts of the song menu screen, or any other parameters, affect collective outcomes, could run these different designs simultaneously in the different worlds. This methodology would ensure that the subjects interacting with each song menu screen are identical and therefore rule out any concern about different subjects in the different experiments. However, we could not use this strategy because we did not know how many subjects we would be able to recruit.

[13]More formally, one way to quantify the amount of social influence on listening decisions would be to calculate the Gini coefficient of the probabilities of listening to songs of different ranks. If there was no social influence all songs would have approximately the same probability of being listened to and the Gini coefficient would be 0; larger Gini coefficients would indicate more inequality in listening probabilities, and thus, more social influence. The Gini coefficient for the listening probabilities in experiment 1 was 0.16, which is substantially lower than the value of 0.42 that was measured in experiment 2. Thus, we have strong evidence that there was greater social influence on listening decisions is experiment 2.



(a) Experiment 1 (b) Experiment 2

Figure 3.18: Probability that a subject in each condition listened to a song of a given market rank in experiments 1 and 2. Subjects' listening decisions were more strongly influenced by market rank in experiment 2 than experiment 1. Unsmoothed results, which are qualitatively the same, are presented in figure B.23.

Another interesting pattern in the listening probabilities in experiment 2, and not present in experiment 1, was the slight increase in propensity for subjects to listen to the least popular songs, those ranked 45th to 48th. This pattern could be an artifact of the list format used in experiment 2 which, in addition to making the top-ranked songs more salient, also made the bottom-ranked songs more salient. Or, instead of being an artifact of our experiment, this could be a real behavioral tendency for some people to want to listen to the least popular songs, perhaps as a form of anti-conformist behavior (Simmel, 1957; Heath et al., 2006). Further experiments would be required to adjudicate between these possibilities.

3.3.2 Collective behavior

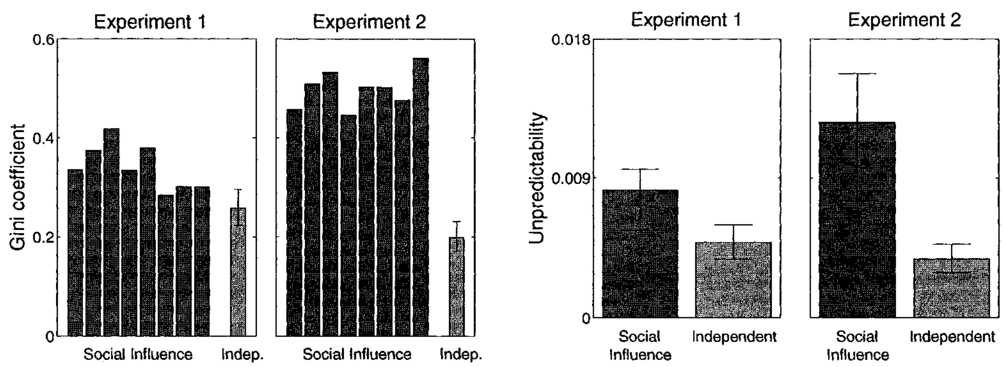
As demonstrated in the previous section, changing the layout of the song menu had the predicted effect at the individual level of increasing the amount of social influence on behavior. The conceptual model suggests that increasing the social influence experienced by individuals would cause increased inequality in success for the songs. Figure 3.19(a) plots the Gini coefficients from experiment 1 and experiment 2. In experiment 2, the Gini coefficients in the influence worlds range from 0.45 to 0.56, with a median value of 0.50 which is substantially larger than the inequality in the influence worlds in experiment 1.[14] In fact, all the social influence worlds in experiment 2 were more unequal than the most unequal world in experiment 1. Using current levels of income inequality as a point of reference, in experiment 1 the inequality in the influence worlds was similar to that in Western Europe ($G \approx 0.33$), while in experiment 2 the inequality was similar to that in developing nations like Nigeria, Peru, and Venezuela ($G \approx 0.50$) (UNDP, 2004). The conceptual model also suggests that increased social influence at the individual level would lead to greater unpredictability at the aggregate level. In figure 3.19(b), we plot the unpredictability measure U (defined in equation 3.4) in experiment 1 and experiment 2 and observe that, as predicted, the unpredictability in the social influence condition increased by about 50% over the values observed in experiment 1.

We can also examine the dynamics of these measures to see if the results might have been different if the experiment had run for a longer (or shorter) period of time. Figures 3.20 and 3.21 suggest that our measure of inequality had largely stabilized, and figure 3.22 suggests that our measure of unpredictability had also stabilized.

3.3.3 Conclusion

Overall, the results from experiment 2 strongly support the proposed relationship between social influence and inequality and unpredictability of success. Since both experiments drew their subjects from the same source and used the same 48 songs, the most likely explanation for the increase in inequality and unpredictability in the social influence condition is the increase in social influence at the individual

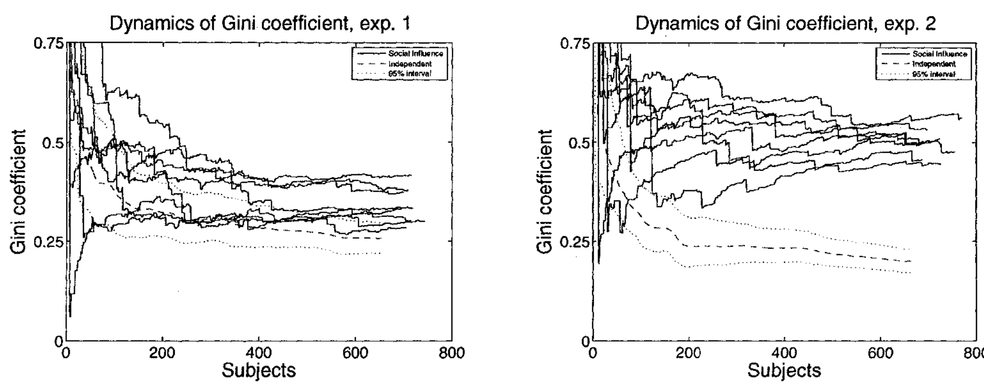
[14]The Gini coefficient in the independent world decreased from experiment 1 to 2 because of a decrease in inequality in the number of song listens per song between the experiments.



(a) Inequality

(b) Unpredictability

Figure 3.19: Plots examining the inequality and unpredictability in experiments 1 and 2. Increasing the amount of social influence increased the median Gini coefficient of the social influence worlds from 0.33 to 0.50. This difference is substantial and is of the same magnitude as the difference between current levels of income inequality in countries like France, the Netherlands, and Spain ($G \approx 0.33$) and countries like Nigeria, Peru, and Venezuela ($G \approx 0.50$) (UNDP, 2004). Increasing the social influence also increased the unpredictability (U) by 50%.



(a) Experiment 1

(b) Experiment 2

Figure 3.20: Dynamics of the Gini coefficient in experiment 1 and 2.

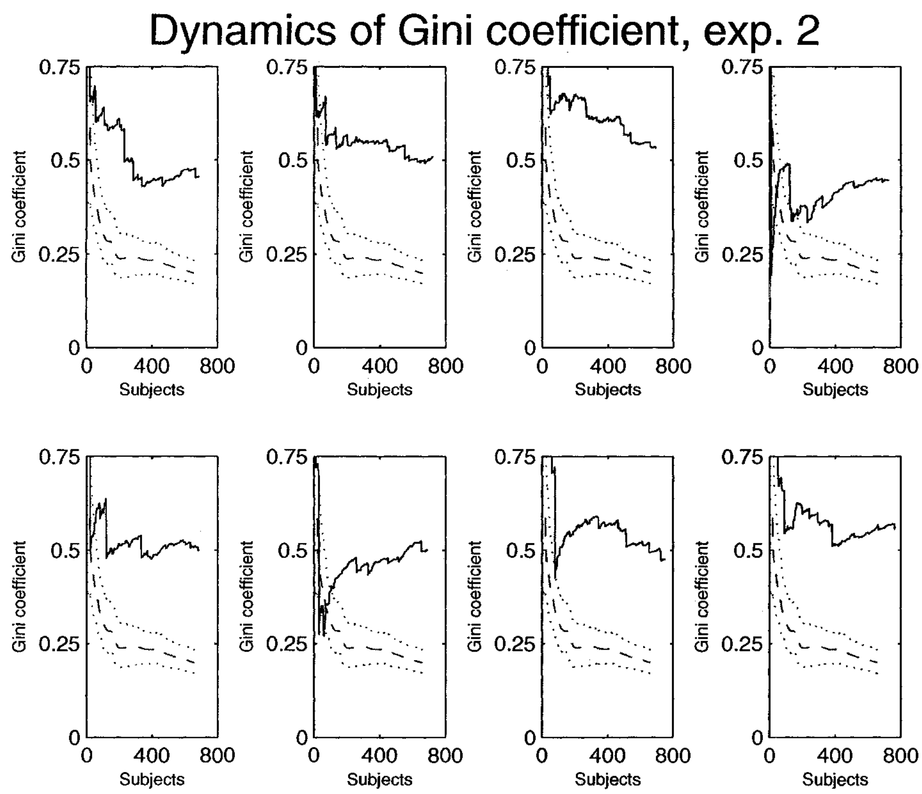
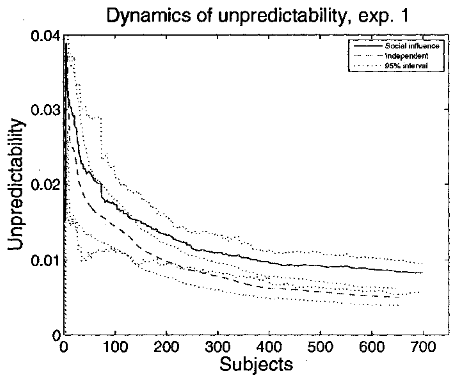
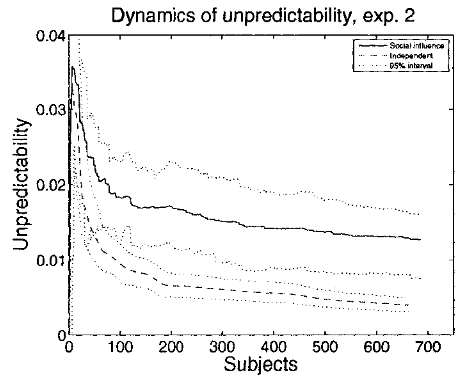


Figure 3.21: Dynamics of the Gini coefficient in experiment 2 for each world. Solid lines are the social influence worlds; the dashed line is the independent world; and the dotted lines represent 95% Monte Carlo intervals around the average independent world.



(a) Experiment 1



(b) Experiment 2

Figure 3.22: Dynamics of unpredictability in experiment 1 and 2.

level.[15] A comparison of experiments 1 and 2 also highlights the surprising effect of social influence. On the one hand, increasing the social influence, by changing the song menu from an unordered grid to a ranked list, increased the inequality of outcomes; thus "the winners" in any particular realization seemed more distinguishable from "the rest." On the other hand, however, this same change also increased differences between the realizations. Increasing social influence, in other words, increased the *appearance* of predictability while simultaneously decreasing the *actual* (i.e., ex-ante) predictability. Critically, because individual participants only ever experience one outcome, the increase in unpredictability is largely invisible to them.

3.4 Appeal and success

The results from these two experiments support our conceptual model without requiring any measure of the "quality" of the songs. However, given the central role that "quality" is thought to play in determining the success of cultural ob-

[15]Experiment 2 also demonstrated that our conceptual model was able to suggest interventions at the individual level which led to predictable effects at the aggregate level. The ability to guide interventions is an accomplishment given that many interventions in social systems have unanticipated (and undesired) consequences (Merton, 1936; Portes, 2000).

jects (Rosen, 1981; Ginsburgh, 2003), further analysis is called for. Previous attempts to measure the "quality" of music, by computer analysis of the acoustic properties of songs (Hamlen, 1991, 1994) or by the ratings of "experts" (Krueger, 2005), have been problematic.[16] In fact, we suspect that any measurement of "quality" will be problematic because no agreed upon definition exists (Gans, 1974; Bourdieu, 1984; DiMaggio, 1987).

Here we avoid the theoretical swamp that is "quality" by instead measuring the *appeal* of each of our songs, which we define to be the independently expressed preferences of a specific population for a specific object. Thus, to measure the appeal of an object, one could imagine something like an opinion poll where a sample of the desired population would be asked to independently express their opinion about the object. The summation of these independently expressed preferences would then provide a measure of the appeal of the object. Because the participants in the independent condition are a random sample of experimental participants (due to random assignment to condition) and because these participants listened to and downloaded songs independently, the market shares in the independent condition provide a natural measure of the appeal of the songs to our population of participants.

Again, we want to emphasize that this measure of appeal is *not* a measure of universal quality, in part because it is specific to a subject pool. Were we to re-run the experiment with a new pool of participants, say senior citizens from Japan, our procedure would almost certainly results in a different measure of appeal. Further, because the measure of appeal is based on market share, it is constrained to sum to

[16]The approach of Hamlen (1991, 1994), was to measure the "quality" of an artist by analyzing the higher frequency harmonics of a sound recording of that artist singing the word "love." However, it is not at all clear that this procedure accurately measures an artist's or song's "quality." More recently, Krueger (2005) attempted to measure the "quality" of a band based on the number of millimeters of print (text and images) devoted to that band in *The Rolling Stone Encyclopedia of Rock & Roll.* Perhaps intuitively more reasonable than the approach by Hamlen, this measurement suffers because it is mostly a measure of success, which, as we shall see, may not be closely related to "quality."

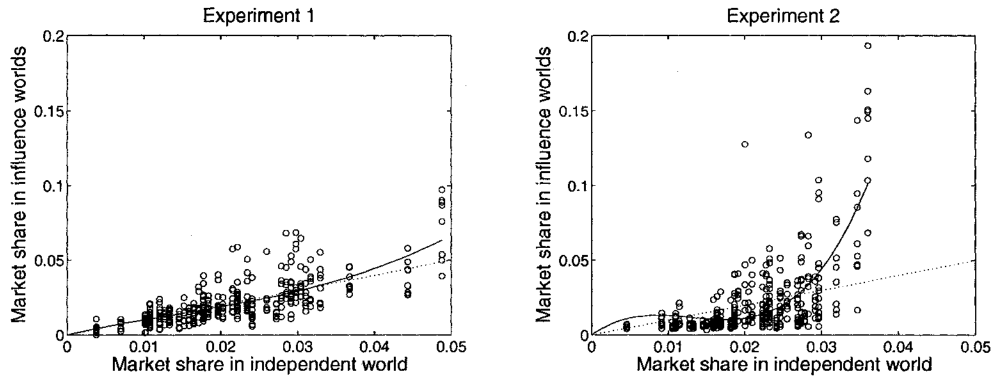
1. Thus, were we to add another song to our experiment, the measured appeal of the other songs would decrease. While these characteristics limit the general applicability of our measure, fortunately, they do not affect the current analysis because we are comparing indistinguishable groups of subjects evaluating the same set of songs.

We now return to a slightly modified version of our original question and ask about the role of appeal in the success of cultural objects. Earlier in this chapter we found that success in the social influence worlds was somewhat unpredictable, but this does not imply that the success of the songs was purely random; instead, success was positively related to appeal. Figure 3.23 plots the relationship between the appeal of the songs, as measured in the independent condition, and the success of the songs in the eight influence worlds. A clear pattern emerges that, in general, songs of higher appeal were more successful; however, this general pattern was quite noisy. Further, while appeal was positively associated with success in both experiments, this relationship was highly nonlinear in experiment 2 as seen in the solid line in the figures which represents the best fit third-degree polynomial to the data. This nonlinearity means that for the best songs in experiment 2, a small increase in appeal was associated with a large increase in success.[17]

Another pattern in figure 3.23 is the cone-shaped spread which implies that the variation in success increased as the appeal increased. Thus, contrary to what one might expect, it was easier to predict the success of low appeal objects than high appeal objects: in other words, it was easier to predict failure than success. Moreover, the unpredictability of these highest appeal objects increased as the amount of social influence at the individual level increased.

In many cultural markets the actual level of success is often not as important as the relative level of success, and therefore an examination of the results in terms

[17]This result is consistent with the convex relation hypothesized in Rosen (1981). However, the origins of convexity in Rosen's model are quite different from those in our experiment where the convexity is a consequence of social influence.



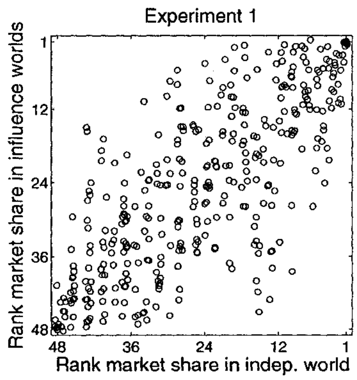
(a) Experiment 1, all influence worlds, market share

(b) Experiment 2, all influence worlds, market share

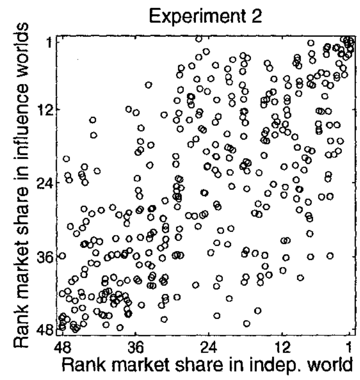
Figure 3.23: Relationship between appeal and success. In experiments 1 and 2 the songs of higher appeal tended to have more success in the social influence worlds, on average. The dotted line indicates the appeal-equals-success line and the solid line is a third-degree polynomial fit to the data. The shape of this polynomial indicates that the relationship between appeal and success is more nonlinear in experiment 2 than experiment 1.

of ranks is natural. Figure 3.24 plots the rank appeal of a song compared with its eight rank success outcomes. As in the analysis of the raw (unranked) data, we see a positive relationship between appeal and success. However, the rank data exhibit tremendous scatter indicating that the relationship between the two measures is weak. An examination of ranked data also reveals a different pattern in terms of unpredictability. In terms of market share, the highest appeal objects were the most unpredictable, but in terms of rank, songs in the middle of the appeal distribution were most unpredictable.[18] Songs in the top quarter of appeal tended to finish in the top half of success and the songs in the bottom quarter of appeal tended to finish in

[18]The differences between the results from the unranked and rank data have two sources. First, the rank of a song can be affected by just a few downloads making rank a partially unstable measure. In contrast, a few downloads cannot substantially change the the market share of a song. The second source of difference is that the rank data collapses much of the variability of results for the top songs. For example, "She Said" by Parker Theory was the highest appeal song in experiment 2. It often finished first in success, but even within the times it finished first, there was a range of different market share values. This variation is lost when the market share values are converted to ranks.



(a) Experiment 1, all influence worlds, rank



(b) Experiment 2, all influence worlds, rank

Figure 3.24: Relationship between rank appeal and rank success in experiments 1 and 2. Songs in the top quarter of appeal tended to finish in the top half of success and songs in the bottom quarter of appeal tended to finish in the bottom half of success. For the songs in the middle of the appeal distribution almost any level of success was possible. Plots are jittered to resolve ties.

the bottom half of success. Thus, the "worst" songs never did very well and the "best" songs never did very badly. But, for the songs in the middle of the appeal ranking, almost any level of success was possible. For example, in experiment 2, "Lockdown" by 52metro was in middle of the appeal distribution, ranked 26th out 48. In one of the influence worlds it finished first in success, and in another world made up of an indistinguishable group participants the very same song finished 40th out of 48 songs.

The relationship between appeal and success can also be quantified using Spearman's rank correlation.[19] As is seen in figure 3.25 (gray bars), the relationship between appeal and success was stronger in each of the eight influence worlds in experiment 1 than in the eight influence worlds in experiment 2, suggesting that social influence weakened the relationship between appeal and success.[20]

---

[19]We use Spearman's rank correlation instead of the more familiar Pearson's correlation because of the nonlinear relationship between appeal and success in experiment 2. Spearman's rank correlation is not affected by this nonlinearity because it is a measure of monotonic association not linear association (Kendall and Gibbons, 1990).

[20]In general, these correlations are actually quite high compared to those often observed in the

In an attempt to remove some of the noise that is present in each success measure, we constructed a measure of average success for each song by averaging its market share across the eight influence worlds ($\bar{m}_i = \frac{\sum_{j=1}^{8} m_{i,j}}{8}$ where $m_{i,j}$ is the market share of song i in world j). The relationship between appeal and this average success measure was stronger than the relationship between appeal and success in any specific realization (figure 3.25 (black bars)). This suggests that each outcome is a combination of signal (based on the appeal of the songs) and noise (based on the effects of social influence). By averaging over many realizations, the noise cancels and we are left with a signal which is related to appeal. We suspect that if we had had 8,000 social influence worlds instead of 8, the rank correlation between appeal and average success would be very close to 1. Thus, in our experiment the intuition that the best songs would become the most successful, is true, *but only on average*; in any particular realization the best songs were not always the most successful.[21]

An important limitation of this analysis is that our measure of appeal, because it is based on the behavior of the people in the independent condition, is unaffected by success in the influence worlds. However, in real cultural markets, appeal is likely to be *endogenous*. For example, Cutting (2003) showed that because of what psychologists call the mere-exposure effect (Bornstein, 1989), the more exposure people have to a specific painting, the more they like it. Therefore, as an object becomes more successful and people are increasingly exposed to it, the appeal of that object increases. This endogenous nature of appeal may have the effect of increasing the observed correlation between appeal and success in real cultural markets. But, rather

social sciences. However, we want to emphasize that because our experiment is not attempting to create a real cultural market, the relative values between experiments are more important than the absolute values.

[21] In some sense this is similar to the idea of "unbiasedness" for a statistical estimator. Unbiased does not mean that the estimate will be equal to, or even close to, the true value. Rather, unbiasedness only implies that, on average, the estimate will equal the true value. Thus, one could loosely say that in our experiment success was an unbiased estimator of appeal, but that each estimate had a large standard error.

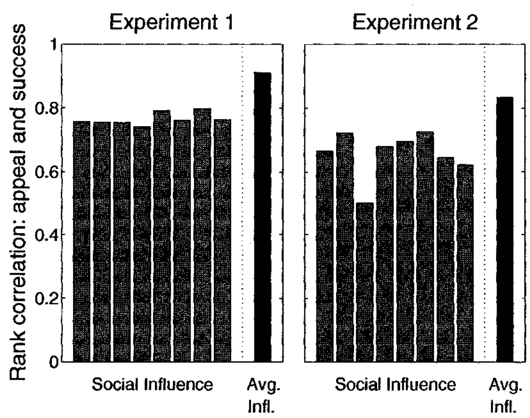


Figure 3.25: Spearman's rank correlation between appeal and success in experiments 1 and 2. Results indicate that increasing social influence from experiment 1 to 2 weakened the relationship between appeal and success. Further, the relationship between appeal and average success was higher than the relationship between appeal and success in any particular realization.

than appeal leading to success, it is probably also the case that success leads to appeal.

3.5 Robustness of results to specific design choices

These two experiments represent only a small portion of the parameter space of all possible experiments using this design. For example, system parameters like the strength and type of social signal, the subject population, the distribution of quality of the songs, and the number of songs probably influence the magnitude of the observed outcomes. Based on our experience with these experiments, we offer a few predictions.

We suspect that other methods of strengthening the social signal would increase the inequality and unpredictability. For example, in our experiment we chose to present the number of previous downloads, the band name, and song name all in the same size font. If, for example, we had presented the download counts in a larger font we suspect that the inequality and unpredictability would be greater.

However, other methods of changing the social signal may have ambiguous effects on outcomes. For example, in our experiments, the social signal was anonymous, in the sense that subjects did not have any information about the characteristics and behavior of previous subjects. If the social signal was instead somehow linked to the identities of the previous subjects, one could imagine that since subjects may be more strongly influenced by "people like them," the cumulative advantage process could be weakened or strengthened depending on the distribution of subjects' identities.

Switching from characteristics of the social signal to characteristics of the songs, we expect that if the songs were more similar in quality, then the inequality in success would be less, but the unpredictability would be greater. Recall, that in these experiments we did not directly set the distribution of quality; rather, it was determined by the songs on `http://www.purevolume.com`. Another key system parameter for the songs is the number used. Because choice overload is so pervasive in cultural markets (Caves, 2000; Vogel, 2004), we chose to use 48 songs in the experiments—the maximum that could fit on a computer screen when presented with the song menu used in experiment 1 (figure 3.2). We conjecture that if we had used more songs, the observed inequality and unpredictability would have increased. Whatever the final number of songs used, it is likely important that this number is much larger than the number of songs that each subject listens to.

Finally, we suspect that the process of social influence observed in the experiments is relatively general, but may be more pronounced with our subject pool (teenagers from the U.S.). We suspect that if the experiment was re-run using a different subject pool that different songs would become successful, but that the overall amount of inequality and unpredictability would be similar. We will address just this question in the next chapter, by replicating experiment 2 with a new pool of participants.

Chapter 4

Experiments 2 and 3

4.1 Introduction

Experiments 1 and 2, presented in the previous chapter, provide strong support for the argument that social influence at the individual-level leads to inequality and unpredictability of success at the collective level. There are risks, however, in drawing sweeping conclusions based solely on studies of American teenagers, a population that may be unusual in many ways.[1] For example, there is some evidence that younger people may be more susceptible to social influence (Park and Lessig, 1977; Pasupathi, 1999) and that conformity in Asch-type experiments varies by country (Bond and Smith, 1996). Therefore, a replication of the results with a different pool of participants would be desirable. In this chapter we present a replication of experiment 2 using a new pool of participants.

In experiment 3, 2,930 participants were recruited by sending an email to

[1]This concern is similar to that shared by psychologists who have questioned what can be learned about human behavior in general from studies of college students (Gordon et al., 1986; Sears, 1986; Peterson, 2001; Mintz et al., 2006). Experimental economists have also recently started asking similar questions (Henrich et al., 2005).

13,546 participants from the electronic small-world experiment (Dodds et al., 2003).[2] This number of participants is smaller than in experiment 2 because experiment 3 had only two social influence worlds and one independent world (as opposed to eight and one) so that we could save as many email addresses as possible for the recruitment of participants for experiment 4 which will be described in the next chapter. This design and sample size again yielded social influence worlds with about 700 participants and an independent condition with about 1,400 participants. As we will see in the next section, participants in experiment 3 differed from those in experiment 2 in their demographic characteristics, behavior, and preferences.

4.2 Comparing participants in experiments 2 and 3

As expected (and hoped for), the participants of experiment 3 differed demographically from those in experiment 2; they were older, more male, and more international (table 4.1). These participants probably differed on unobservable features as well, particularly willingness to be in experiments.[3] Recall that most participants from experiments 1 and 2 were recruited through ads, while most participants from experiment 3 (and later 4) were people who had already participated in one experiment, the electronic small-world experiment, and were willing to participate in another after prompting from a short email. This difference in willingness to participate likely explains a major difference in behavior between the participants: even thought they liked the songs less, as judged by their lower average ratings and their

[2]Because these emails generated addition web-postings, not all participants in experiment 3 had been in the small-world experiment. The text of the recruitment email is presented in section C.2.

[3]Anecdotally, the emails that we received during experiment 3, which tended to offer advice on our research design, were quite different from those received during experiment 2, which tended to complain about the music in the experiment.

Category	`www.bolt.com` Experiment 2 ($n = 7,192$) (% of participants)	Small-world experiment Experiment 3 ($n = 2,930$) (% of participants)
Female	73.9	38.0
Broadband connection	69.0	90.6
Has downloaded music from other sites	62.4	69.3
Country of Residence		
United States	81.8	68.4
Canada	4.4	6.3
United Kingdom	4.7	6.6
Other	9.1	18.7
Age		
14 and younger	16.0	1.5
15 to 17	34.9	5.7
18 to 24	39.2	29.8
25 and older	9.9	63.1

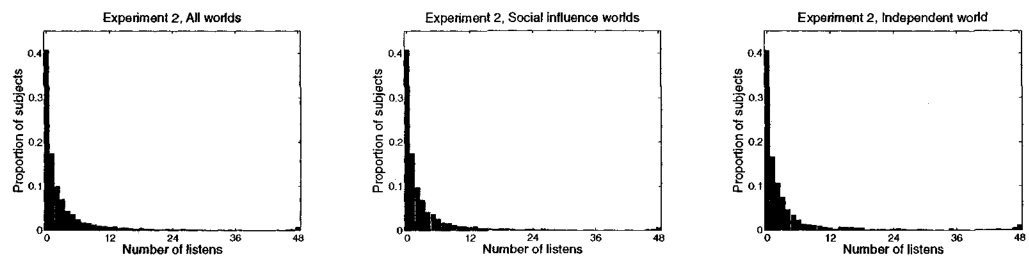
Table 4.1: Descriptive statistics about the participants in experiments 2 and 3. Most participants from experiment 2 were recruited from `www.bolt.com`. Most participants from experiment 3 were recruited by emails to participants in the electronic small-world experiment (Dodds, Muhamad, and Watts, 2003) and the subsequent web postings these emails generated. Participants in experiment 3 were older, more male, and more international.

lower probability of download given listen, participants in experiment 3, on average, listened to many more songs, 7.7 compared to 3.6 (table 4.2). Figures 4.1, 4.2, 4.3, and 4.4 plot the distribution of number of listens and downloads per participant in experiments 2 and 3.

In addition to liking the songs less overall, participants in experiment 3 liked somewhat different songs as can be seen by comparing the measured appeal in experiments 2 and 3 (recall, that appeal is defined as the market share of downloads in the independent world). Figure 4.5(a) shows that there was a positive relationship between appeal in experiment 2 and 3, but the relationship was not particularly

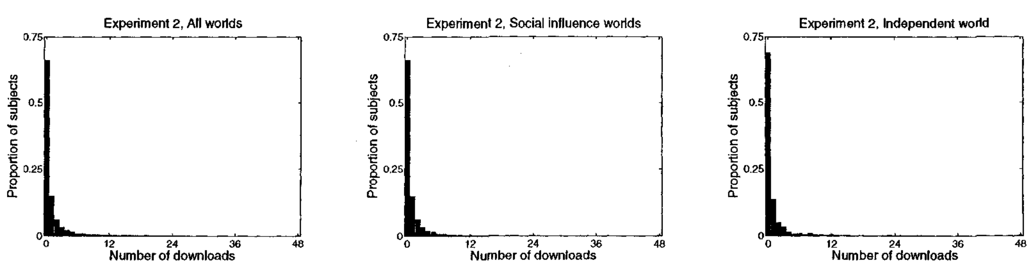
	Experiment 2			Experiment 3		
	Social influence 8 worlds ($n = 5,746$)	Independent 1 world ($n = 1.446$)	Total 9 worlds ($n = 7,192$)	Social influence 2 worlds ($n = 1,471$)	Independent 1 world ($n = 1,459$)	Total 3 worlds ($n = 2,930$)
Number of listens	20,217	5,643	25,860	10,591	11,844	22,435
Mean per participant	3.5	3.9	3.6	7.2	8.1	7.7
Number of downloads	8,106	2,192	10,298	2,040	1,691	3,731
Mean per participant	1.4	1.5	1.4	1.4	1.2	1.3
Pr[download | listen]	0.401	0.377	0.398	0.193	0.143	0.166
Mean rating (# of stars)	3.31	3.22	3.29	2.70	2.55	2.62

Table 4.2: Descriptive statistics of subject behavior in experiment 2 and 3.



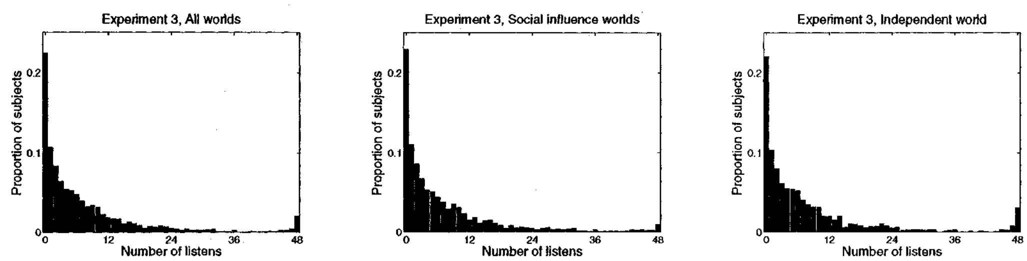
(a) Exp 2, all worlds (b) Exp 2, social influence worlds (c) Exp 2, independent world

Figure 4.1: Distribution of the number of listens per participant in experiment 2. Results are also presented for just participants in the social influence worlds and just participants in the independent world.



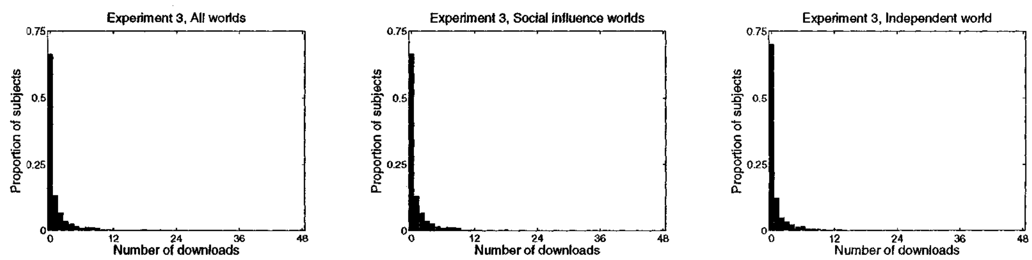
(a) Exp 2, all worlds (b) Exp 2, social influence worlds (c) Exp 2, independent world

Figure 4.2: Distribution of the number of downloads per participant in experiment 2. Results are also presented for just participants in the social influence worlds and just participants in the independent world.



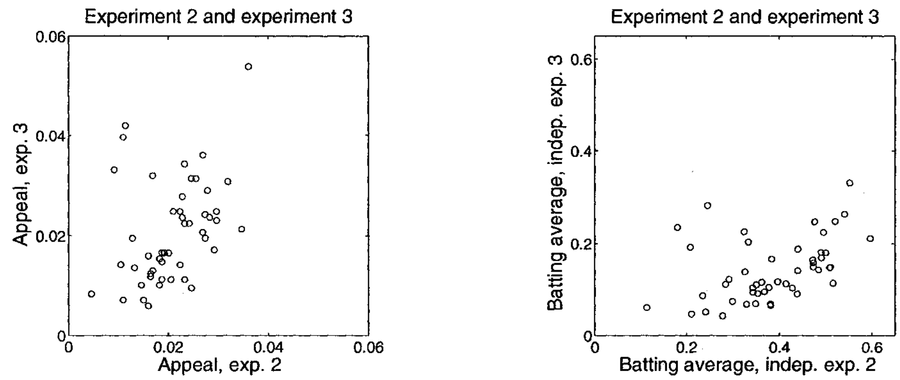
(a) Exp 3, all worlds (b) Exp 3, social influence worlds (c) Exp 3, independent world

Figure 4.3: Distribution of the number of listens per participant in experiment 3. Results are also presented for just participants in the social influence worlds and just participants in the independent world.



(a) Exp 3, all worlds (b) Exp 3, social influence worlds (c) Exp 3, independent world

Figure 4.4: Distribution of the number of downloads per participant in experiment 3. Results are also presented for just participants in the social influence worlds and just participants in the independent world.



(a) Comparing market share

(b) Comparing batting average

Figure 4.5: Comparing the market shares and batting averages in the independent condition in experiment 2 and 3. The rank correlation between the market share in the independent worlds in experiment 2 and experiment 3 was 0.42.

strong ($\rho = 0.42$).[4] Another way to compare the preferences of the participants is to compare the batting averages in the independent conditions. Figure 4.5(b) shows that, in general, the batting averages were much lower in experiment 3.

In review, participants in experiment 3 differed from those in experiment 2. They were older, more male, and more international, more cooperative, and had different musical preferences. In the next section we will explore whether, despite these differences, experiment 3 replicated the results of experiment 2.

4.3 Results from experiment 3

Just as in experiments 1 and 2, we can examine the effect of the social influence on subjects' decisions about which songs to listen to. At the time each subject participated, every song in their world had a specific market share and market rank

[4]The song that improved the most in appeal from experiment 2 to experiment 3 was "Gnaw" by Silverfox, an easy-listening electronic piece. In contrast, the song that decreased the most in appeal from experiment 2 to 3 was "The End in Friend" by The Broken Promise. This hard-driving, angsty song includes the chorus: "Tie the noose around my neck \ Kick that chair beneath my feet \ Watch me choking red \ As you turn and walk away."

(for example, the song with the highest market share has a market rank of 1). We can measure the influence of market rank on subjects' listening decisions by calculating the probability that a subject in the social influence condition chose to listen to the song of a given market rank (independent of which song occupied that rank at the time). For example, figure 4.6 shows that in the influence worlds each subject had about a 65% chance of listening to whatever was the most popular song, about a 10% chance of listening to whatever was of middle popularity, and about a 30% chance of listening to whatever was the least popular.[5] This preference for listening to the least popular songs—which also occurred in experiment 2—could be an artifact of the list format used which, in addition to making the top-ranked songs more salient, also made the bottom-ranked songs more salient. Or, instead of being an artifact of our experiment, this could be a real behavioral tendency for anti-conformist behavior (Heath et al., 2006). Whatever the reason, however, the listening pattern in the social influence condition was strikingly different from that in the independent condition where the popularity of the song, which was not available to the subjects, had no affect on listening decisions.

Given that social influence occurred at the individual level we can turn our attention to the collective-level outcomes of inequality and unpredictability. Figure 4.7(a) shows that the Gini coefficient—described at length in the previous chapter—was greater in the social influence worlds than the independent world. Further, figure 4.7(b) shows that the unpredictability—again described at length in the previous chapter—was greater in the social influence worlds than the independent world. Thus, both of our main hypotheses held in this different study population suggesting that the effects of social influence that were observed in experiments 1 and 2 cannot be

[5]The values presented in figure 4.6 have been smoothed to aid visualization. The unsmoothed results, which are qualitatively similar, are presented in figure C.7(b). Also, the probabilities in figure 4.6 do not sum to 1 because the figure has 95 data points: the 48 possible ranks plus the 47 possible ties (e.g., two songs that were tied in first place were both be assigned a rank of 1.5).

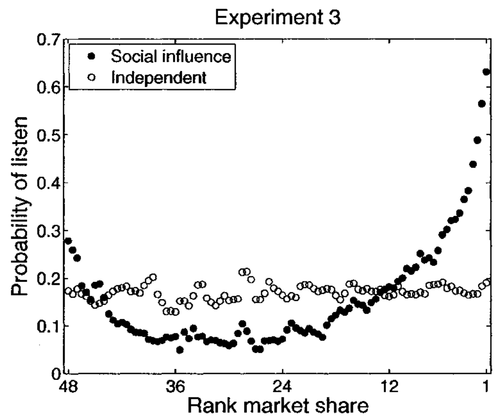
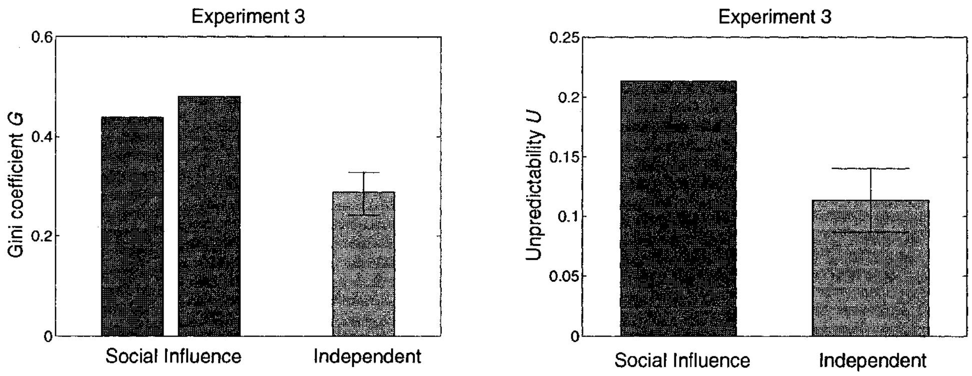


Figure 4.6: Probability that a subject in each condition listened to a song of a given market rank in experiment 3. Unsmoothed results, which are qualitatively the same, are presented in figure C.7

attributed to some special characteristics of the teenagers from `www.bolt.com`. In the next section, we will undertake a closer comparison of the results experiments 2 and 3.

4.4 Comparison of results from experiments 2 and 3

At first glance it appears that experiment 3 provided a complete replication of experiment 2. However, further comparison between the results of the two experiments yields some important differences. First, the inequality in the social influence worlds was smaller in experiment 3, while the inequality in the independent world was larger (figure 4.8(a)). Also, the unpredictability in the social influence condition was smaller, while the unpredictability in the independent condition was larger (figure 4.8(b)). In other words, in experiment 3 the differences in inequality and unpredictability across conditions were smaller than the differences in experiment 2



(a) Inequality

(b) Unpredictability

Figure 4.7: Inequality and unpredictability in experiment 3. There are no error bars around the unpredictability in the social influence condition because there were only two such worlds.

(figure 4.8).

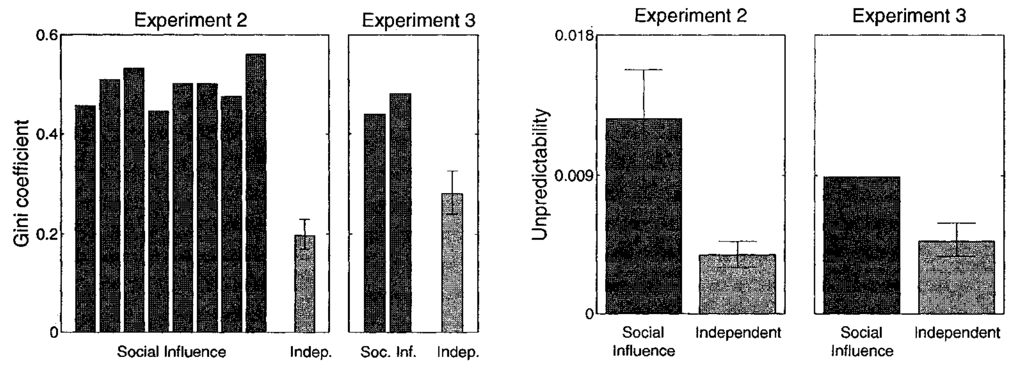
In addition to comparing the final level of inequality and unpredictability, we can also compare the dynamics of these measures in the two experiments, plotted in figures 4.9 and 4.10. Overall these figures are quite similar in shape and seem to indicate that we would not have observed substantially different results if either experiment had run for longer.

We call also compare the individual level behavior across the two experiments (figure 4.11). Qualitatively the listen choice plots are the same, but future work could go into attempting to quantify the differences.

To conclude, experiment 3 provided a replication of the findings of experiment 2, but there were also some differences between the two. Unfortunately, our model, which is theoretical, as opposed to mathematical, cannot explain these differences. Experiment 3, therefore, provides further motivation for developing a more detailed mathematical model of this process, as well as a set of facts which much be reproducible by such a model.[6]

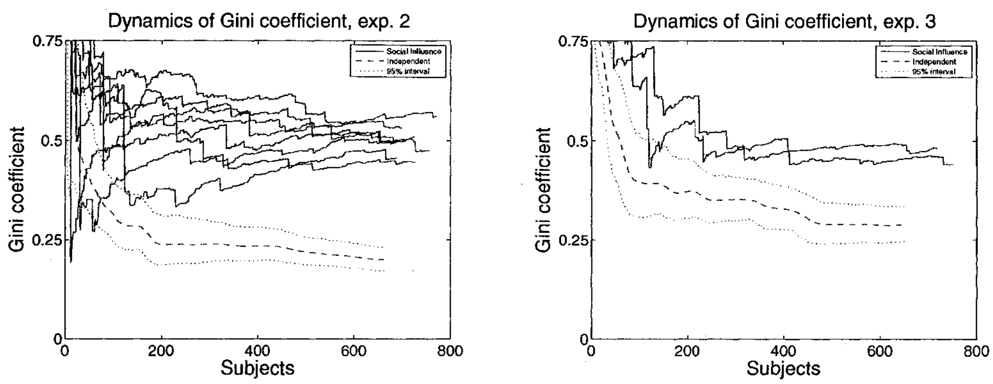
---

[6]We suspect that these differences are because in experiment 3 participants listened to more



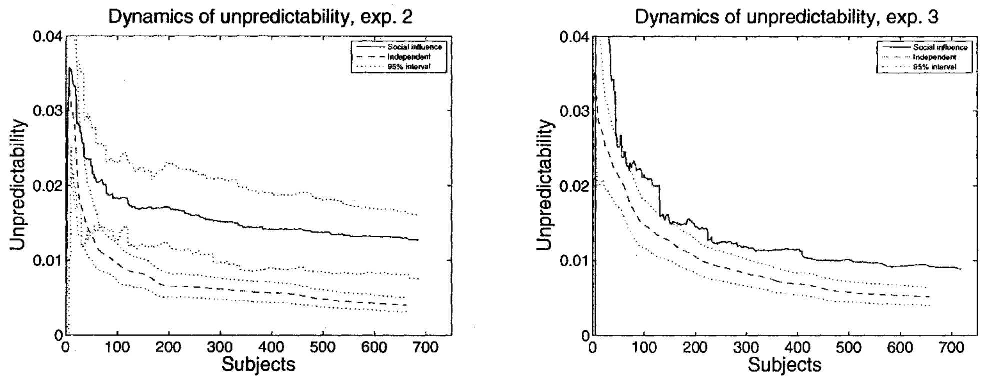
(a) Inequality (b) Unpredictability

Figure 4.8: Comparing the inequality and unpredictability in experiment 2 and 3. There are no error bars around the unpredictability in the social influence condition in experiment 3 because there were only two such worlds.



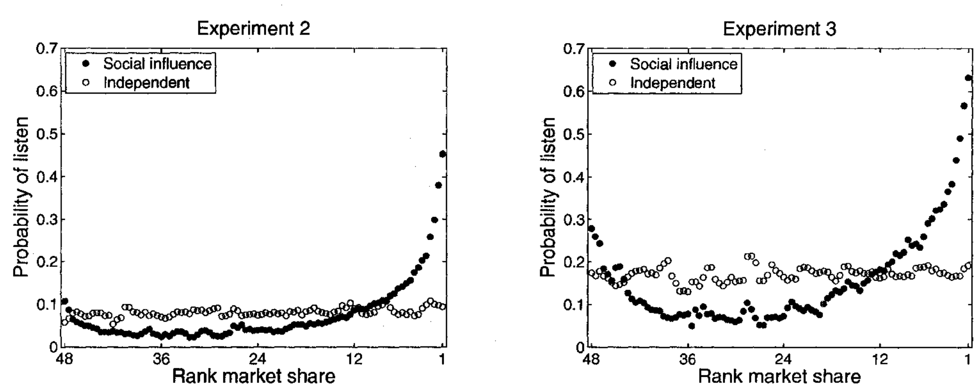
(a) Experiment 2 (b) Experiment 3

Figure 4.9: Dynamics of the Gini coefficient in experiments 2 and 3.



(a) Experiment 2 (b) Experiment 3

Figure 4.10: Dynamics of unpredictability in experiments 2 and 3.



(a) Experiment 2 (b) Experiment 3

Figure 4.11: Probability that a subject in each condition listened to a song of a given market rank. Unsmoothed results, which are qualitatively the same, are presented in figure C.7.

4.5 Cross-experiment prediction

The results of experiments 2 and 3 also allow us to explore the possibility of predicting success outcomes from a variety of different data sources. Mapping this question onto a practical situation, we can imagine the independent world as being similar to a focus group and the social influence worlds as being similar to a market outcome. Three questions then become:

1. How well can you predict success with the "right" participants in the focus group (i.e., using the focus group made up of experiment 2 participants to predict experiment 2 success)?

2. How well can you predict success with the "wrong" participants in the focus group (i.e., using the focus group made up of experiment 2 participants to predict experiment 3 success)?

3. How well can you use market success in one context to predict market success in another context (i.e., success in experiment 2 to predict success in experiment 3)?

In the next sections we will address these questions in a variety of ways, but the reader should keep in mind that we cannot answer these question "in general." Rather, we can just present the results from our experiment. Readers, therefore, should exercise caution when attempting to generalize these results to other situations.[7]

---

songs on average than in experiment 2: 7.7 vs 3.6 out of 48 possible songs. We have no evidence, however, to support this suspicion.

[7] A fourth question, that of using market success to predict market success in the same context is not directly explored because it does not map onto a real-world situation. However, the data to examine this question is presented in the appropriate tables and graphs.

4.5.1 Prediction of market share

In order to define how well a given piece of information allows for the prediction of another piece of information we must have some prediction algorithm which turns the available information into a prediction. The most obvious, and possibly most natural, way to predict market share of all songs in world j, given the market share of all songs in world i, is simply to predict that the market shares will be the same in both worlds,

$$\widehat{m_{sj}} = m_{si} \tag{4.1}$$

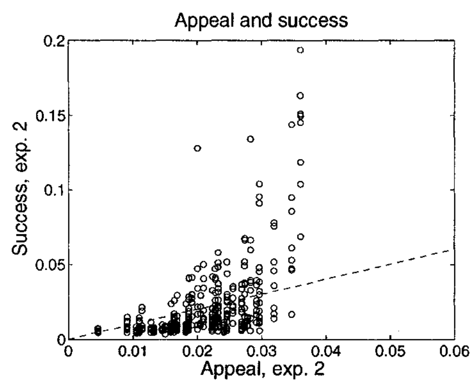
where m_{si} is the market share of song s is world i. Using the data from our experiments and this algorithm we can attempt to address our three motivating questions.[8]

To develop some intuition about the likely performance of our prediction algorithm in these situations, figure 4.12 plots the success measured in experiments 2 and 3 as a function of appeal measured in experiments 2 and 3. Figures 4.12(a) and 4.12(d) show that even when using appeal to predict success within an experiment there is likely to be substantial error. Further, figures 4.12(b) and 4.12(c) show that when using appeal to predict success across experiments there appears to be even more error. Regarding the third question of using success in one context to predict success in another context, figure 4.13 presents a scatter plot of success in experiment 3 as a function of success in experiment 2.[9] However, just like the previous plots, this one seems to suggest there will be substantial prediction error.

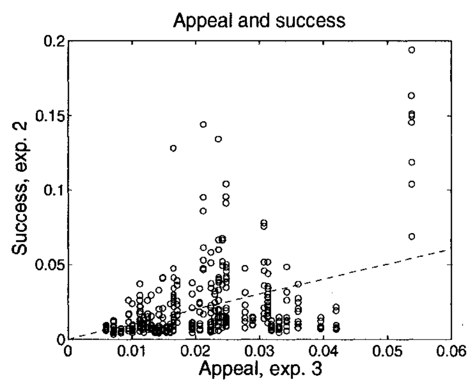
While these scatter plots are helpful for building intuition, it is desirable to quantify the prediction errors by calculating the total song-wise difference in market

[8]Note that this algorithm may not be the ideal algorithm, but it is impossible to address the three questions in section 4.5 in any general way. They can only be evaluated in the context of a specific algorithm.

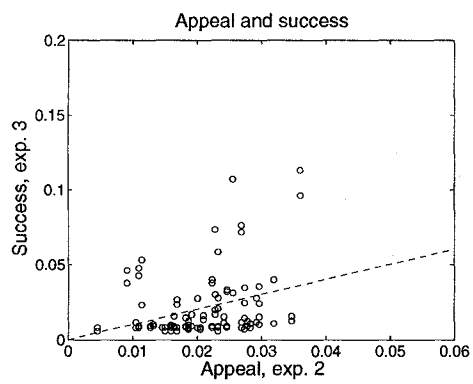
[9]Care should be taken interpreting figure 4.13 because each pair of songs is represented by 16 data points (8 experiment 2 success outcomes $\times$ 2 experiment 3 success outcomes).



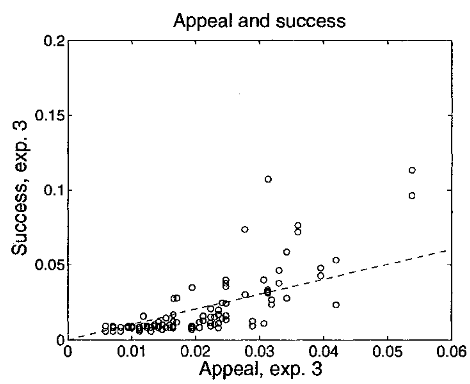
(a) Exp 2 to predict Exp 2



(b) Exp 3 to predict Exp 2



(c) Exp 2 to predict Exp 3



(d) Exp 3 to predict Exp 3

Figure 4.12: Success measured in experiments 2 and 3 as a function of appeal measured in experiments 2 and 3. Prediction of success in experiment 2 with appeal in experiment 3 gives a funny inverse U-shaped relationship, except for the highest appeal song.

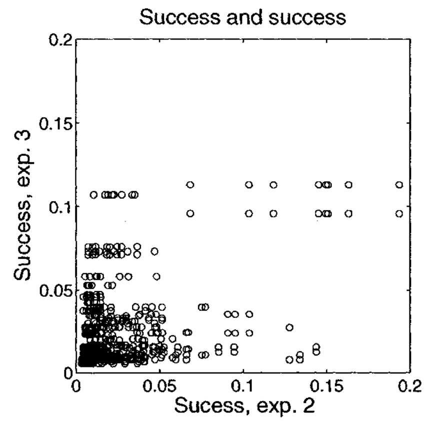


Figure 4.13: Comparing market success in experiment 2 and 3. This plot is somewhat hard to interpret because each pair of songs is represented by 16 data points (2×8).

share between the predicted and actual values,[10]

$$E(m)_{ij} = \sum_{s=1}^{48} \mid \widehat{m_{sj}} - m_{sj} \mid \tag{4.2}$$

where $\widehat{m_{sj}}$ is the predicted market share for song s in world j (defined in equation 4.1) and m_{sj} is the actual market share. $E(m)_{ij}$ is hard to interpret directly, but it ranges from from a minimum of 0, if every song has the same market share in both worlds, to a maximum of 2, if all the market share in world i belongs song s and all the market share in world j belongs to a different song. Further, $E(m)$ is the average of $E(m)_{ij}$ over all relevant worlds,

$$E(m) = \frac{\sum_{ij} E(m)_{ij}}{i \cdot j} \tag{4.3}$$

Table 4.3 presents the errors made by our prediction method across conditions and experiments; higher values of $E(m)$ indicate worse predictions.[11] For example, table 4.3 shows that if you are predicting market share in a social influence world

[10]We could also penalize the errors by the sum of the square differences, but this approach seems to add complication without insight.

[11]Note that the table is symmetric because the definition of error (equation 4.2) does not distinguish predicting world i from j and predicting j from i.

	Experiment 2		Experiment 3	
	Independent	Social influence	Independent	Social influence
Independent (exp 2)	0	0.65	0.35	0.69
Social influence (exp 2)	0.65	0.46	0.70	0.75
Independent (exp 3)	0.35	0.70	0	0.46
Social influence (exp 3)	0.70	0.75	0.46	0.43

Table 4.3: Error in predicting market share, $E(m)$, in experiments 2 and 3. Note that the table is symmetric. The error of predicting an independent condition with itself is 0 because we did not do splits as was done in as in Salganik et al. (2006).

in experiment 2 you are better off having a randomly chosen influence world from experiment 2 than the independent world from experiment 2 (i.e., $0.46 < 0.65$).[12]

Unfortunately, the interpretation of table 4.3 is difficult because $E(m)$ does not have a natural scale or meaning. To improve understanding we can give these results a proportional reduction in error interpretation by comparing the prediction using our method and a given piece of information to the prediction that would be made using no information (Costner, 1965). In this case the prediction with no information would be that each song will earn the average market share $\frac{1}{48}$,

$$E(m)_i^{naive} = \sum_{s=1}^{48} \mid \frac{1}{48} - m_{si} \mid \tag{4.4}$$

This naive prediction can then be average over all worlds in a condition to give,

$$E(m)^{naive} = \frac{\sum_i E(m)_i^{naive}}{i} \tag{4.5}$$

The error from the naive predictions are presented in table 4.4.[13]

By comparing the errors with different prediction methods we can give our predictor (equation 4.1) a proportional reduction in error interpretation (Costner,

[12]This initially surprising results is partially a results of our simple prediction algorithm (equation 4.1). Because the market share in the independent condition is much less skewed, it is not very good at directly predicting the market share in a social influence world. We will turn our attention to predicting rank shortly (section 4.5.2).

[13]The size of $E(m)^{naive}$ is determined by the variation in market share.

	Experiment 2		Experiment 3	
	Independent	Social influence	Independent	Social influence
Naive prediction	0.27	0.77	0.40	0.73

Table 4.4: Naive error, $E(m)^{naive}$, in predicting market share in experiments 2 and 3. The reason that the error of the naive estimator is greater in experiment 3 independent than experiment 2 independent is because the variation of outcomes in the independent condition is greater in 3 than 2.

1965):

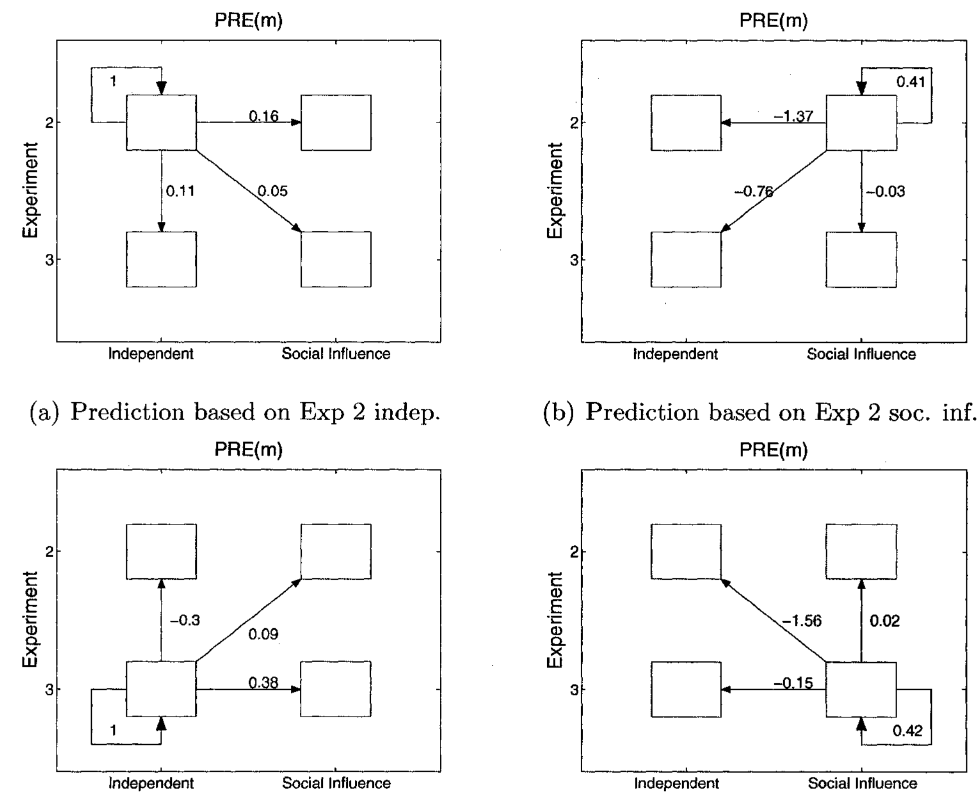
$$PRE(m) = \frac{E(m)^{naive} - E(m)}{E(m)^{naive}} \tag{4.6}$$

where $E(m)$ and $E(m)^{naive}$ are defined in equations 4.3 and 4.5. This measure varies from $-\infty$ to 1, where a negative value means that naive predictions are better than our predictions and a value of 1 means that our predictions have no error.[14] The measure is further illustrated with an example. Let's say you want to predict market share in a social influence world in experiment 2. If you use the naive prediction, your error will be 0.77 (from table 4.4). If you have another social influence world from experiment 2 and you use the predictor in equation 4.1, then your expected error is 0.46 (from table 4.3). Therefore, given our prediction algorithm, having that information reduces your error by 40% ($\frac{0.77-0.46}{0.77}$). The remain 60% of the error is still unexplained. If, instead, you had the outcome from a social influence world in experiment 3 your expected error would be 0.75 and so your reduction in error is about 0 ($\frac{0.77-0.75}{0.77}$). In other words, if you are using our prediction algorithm to predict a result from a social influence world in experiment 2 having a social influence outcome from experiment 3 gives you almost no information—you could just as well guess randomly. These results from all cross comparisons are presented in table 4.5 and visually in figures 4.14.

[14]This proportion reduction in error is similar in spirit to the R^2 measure in linear regression which compares residuals from a full model to residuals from a naive model; for more on this similarity, see Kviz (1981). One differences however is that R^2 varies from 0 to 1 while $PRE(m)$

Proportional reduction in error, $PRE(m)$				
	Experiment 2		Experiment 3	
	Independent	Social influence	Independent	Social influence
Independent (exp 2)	1	0.16	0.11	0.05
Social influence (exp 2)	-1.37	0.41	-0.76	-0.03
Independent (exp 3)	-0.30	0.09	1	0.38
Social influence (exp 3)	-1.56	0.02	-0.15	0.42

Table 4.5: Proportional reduction in error in predicting market share, $PRE(m)$, in experiments 2 and 3.



(a) Prediction based on Exp 2 indep.

(b) Prediction based on Exp 2 soc. inf.

(c) Prediction based on Exp 3 indep.

(d) Prediction based on Exp 3 soc. inf.

Figure 4.14: Proportional reduction in error for predicting market share, $PRE(m)$, in experiments 2 and 3.

Returning to the questions motivating these calculations, using the independent condition to predict success in the social influence condition leads to some reduction in error, and the results are better within experiment 3: $PRE(m)$ of 0.38 compared to 0.16 in experiment 2.[15] In other words, in these set of experiments knowing the results of the "right" focus group will reduce error in prediction, but not substantially.[16] Using the independent condition in either experiment to predict success in the other experiment has almost no value: $PRE(m)$ of 0.05 (exp 2 to exp 3, figure 4.14(a)) and 0.09 (exp 3 to exp 2, figure 4.14(c)). Recasting this result in the framework of focus groups means that a focus group with the "wrong" population leads to almost no improvement in prediction. Finally, if one attempts to use the success outcome from one experiment to predict the success outcome of the other—as for example using the success of a TV show in one country to predict its success in another—again, you get almost no reduction in error over the naive prediction: $PRE(m)$ of -0.03 and 0.02. Thus, in comparing experiment 2 and 3 we can see that the only information which improves prediction of market share in the social influence worlds is the market share in the independent world in that same experiment.

Given these results, we would like to reiterate two important caveats. First, these results are merely descriptive of what we observed in this set of experiments; results in other contexts will probably be different. Second, these results only apply to the prediction algorithm that we used (equation 4.1). In an attempt to present some more robust results, we will next turn our attention to the prediction of market rank, a situation where the appropriate prediction algorithm is clear.

---

varies from $-\infty$ to 1.

[15]This finding is consistent with, but not equivalent to, the decrease in unpredictability between experiment 2 and 3 as seen in figure 4.8(b).

[16]This estimate of reduction in error is probably an underestimate of the benefit of predicting market share because our prediction algorithm is rather simple. A more sophisticated algorithm would know that success in the social influence worlds is more skewed than in the independent world and would therefore account for this; our algorithm does not.

4.5.2 Prediction of market rank

In this section, we attempt to address our three guiding questions, but instead of predicting market share, we attempt to predict market rank. In this case, the most natural prediction algorithm given the outcome in some world i is to predict the same rank in world j,

$$\widehat{k_{sj}} = k_{si} \tag{4.7}$$

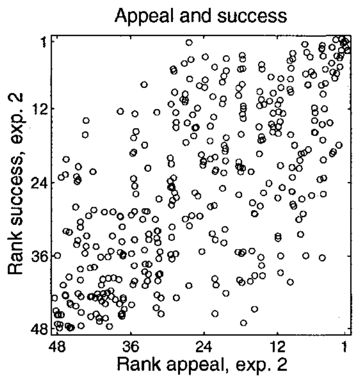
where k_{si} is the market rank of song s is world i. To build intuition, figure 4.15 plots the market rank measured in experiments 2 and 3 as a function of rank appeal measured in experiments 2 and 3. Figures 4.15(a) and 4.15(d) show that even when using the rank appeal to predict market rank within an experiment there is going to be considerable error, particularly for songs in the middle of appeal distribution. When trying to predict market rank across experiments, however, there appears to be even more error (figures 4.15(c) and 4.15(b)). Regarding the third question of using success in one context to predict success in another context, figure 4.16 presents a scatter plot of market rank in experiment 3 as a function of market rank in experiment 2.[17] However, just like the previous plots, this one seems to suggest there will be substantial prediction error.

As in the previous section (4.5.1), we can quantify the magnitude of the errors in prediction of market rank by calculating the total song-wise difference in rank between two randomly chosen realizations,

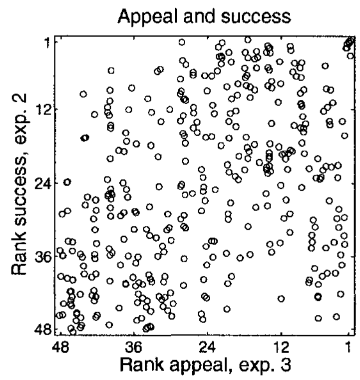
$$E(k)_{ij} = \sum_{s=1}^{48} \mid k_{si} - k_{sj} \mid \tag{4.8}$$

where i is the world that you have information about, j is the world that you are trying to predict, and k_{si} is the market rank of song s in world i. Results are

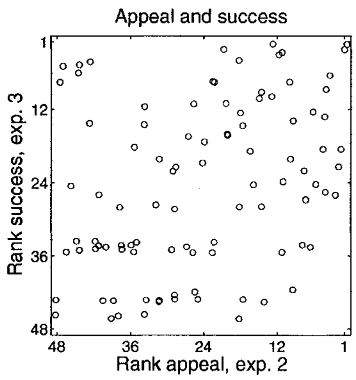
[17] Just as with figure 4.13, care should be taken interpreting figure 4.16 because each pair of songs is represented by 16 data points (8 experiment 2 success outcomes $\times$ 2 experiment 3 success outcomes).



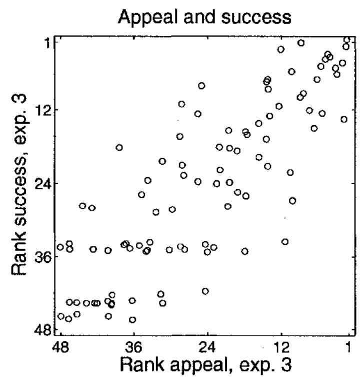
(a) Exp 2 to predict Exp 2



(b) Exp 3 to predict Exp 2



(c) Exp 2 to predict Exp 3



(d) Exp 3 to predict Exp 3

Figure 4.15: Market rank measured in experiments 2 and 3 as a function of rank appeal measured in experiments 2 and 3.

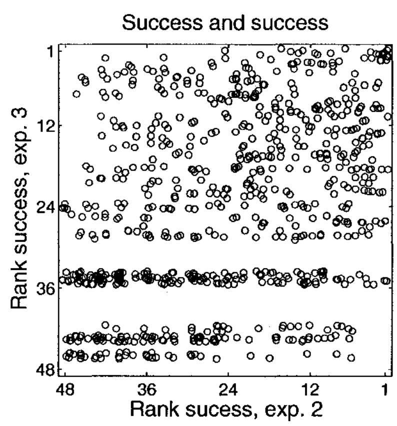


Figure 4.16: Comparing market rank in experiment 2 and 3. This plot is somewhat hard to interpret because each pair of songs is represented by 16 data points (2×8).

	Experiment 2		Experiment 3	
	Independent	Social influence	Independent	Social influence
Independent (exp 2)	0	415.5	518	603
Social influence (exp 2)	415.5	471	569.75	587
Independent (exp 3)	518	569.75	0	288
Social influence (exp 3)	603	587	288	382

Table 4.6: Error in predicting market rank, $E(k)$, in experiments 2 and 3. Note that the table is symmetric. Also the error of predicting an independent condition with itself is 0 because we did not do splits as was done in as in Salganik et al. (2006).

Proportional reduction in error, $PRE(k)$

	Experiment 2		Experiment 3	
	Independent	Social influence	Independent	Social influence
Independent (exp 2)	1	0.28	0.10	-0.05
Social influence (exp 2)	0.28	0.18	0.01	-0.02
Independent (exp 3)	0.10	0.01	1	0.50
Social influence (exp 3)	-0.05	-0.02	0.50	0.34

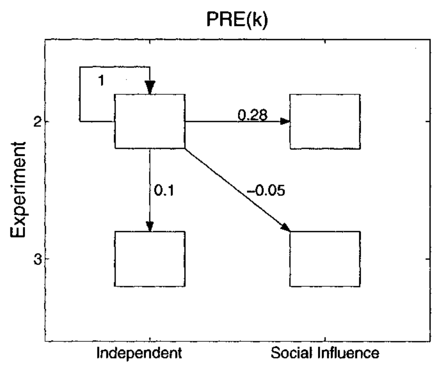
Table 4.7: Proportion reduction in error in predicting market rank, $PRE(k)$, in experiments 2 and 3.

presented in table 4.6. Again, these results are difficult to interpret directly so we re-express them in terms of proportion reduction in error in table 4.7, where,

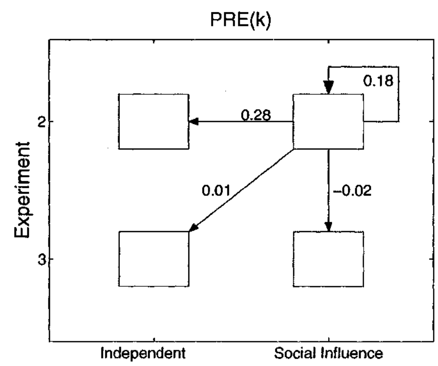
$$PRE(k) = \frac{E(k)^{naive} - E(k)}{E(k)^{naive}} \tag{4.9}$$

with $E(k)_i^{naive} = \sum_{s=1}^{48} \mid 24.5 - k_{si} \mid$, the naive prediction of rank. $E(k)$ is then the average of $E(k)_{ij}$ over all appropriate worlds.

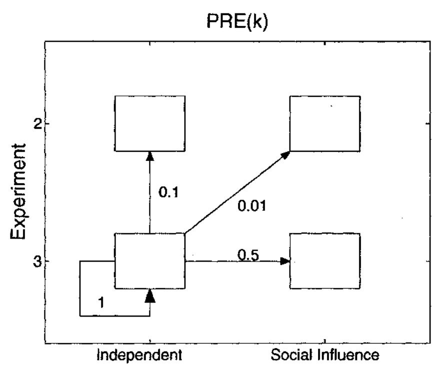
Returning to the questions motivating these calculations, using rank appeal to predict market rank leads to some reduction in error, and the results are better within experiment 3: $PRE(k)$ of 0.50 compared to 0.28 in experiment 2. In other words, in these set of experiments using the "right" focus group will reduce error in prediction of market rank. Using rank appeal in either experiment to predict market rank in the other experiment has almost no value: $PRE(k)$ of -0.05 (exp 2 to exp 3, figure 4.17(a)) and 0.01 (exp 3 to exp 2, figure 4.17(c)). Recasting this



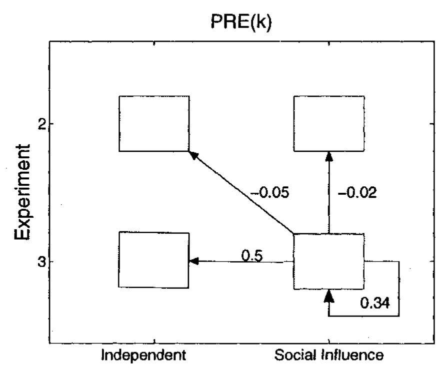
(a) Prediction based on Exp 2 indep.



(b) Prediction based on Exp 2 soc. inf.



(c) Prediction based on Exp 3 indep.



(d) Prediction based on Exp 3 soc. inf.

Figure 4.17: Proportional reduction in error for predicting market rank, $PRE(k)$, in experiments 2 and 3.

result in the framework of focus groups means that a focus group with the "wrong" population leads to almost no improvement in prediction of market rank. Finally, if one attempts to use the market rank from one experiment to predict the market rank in the other—as for example using the success of a TV show in one country to predict its success in another—again, you get almost no reduction in error over the naive prediction: $PRE(k)$ of -0.02 and -0.02. Thus, as with the prediction of market share, the only information which improves prediction of the social influence worlds is the independent world in that same experiment. Anything other than the "right" focus group has almost no value.

4.5.3 Pairwise prediction

In addition to trying to predict the market share or market rank of the entire set of songs it is also occasionally of interest to predict the relative market share or relative market rank of two songs. For example, we can measure the probability of correctly predicting the ordering of success of two songs in one outcome given their ordering in some other outcome.[18] Results are presented in table 4.8. The probabilities should be compared to 0.5, the probability of correctly guessing the rank ordering with no information. For example, let's say that we are attempting to predict the rank ordering of two songs in a social influence world in experiment 2. If we had the rank ordering of the songs in the independent world in experiment 2 our error would be 0.24 (1-0.76). Therefore, having this information reduces our error by about 50% $\frac{0.5-0.24}{0.5}$. Having information about their rank order in the independent condition in experiment 3 would also reduce our error by a smaller amount, 0.38.

Two surprising results from table 4.8 are that when trying to predict the rank ordering in a social influence world experiment 3 it is better to have the outcome of a

[18]This measure is closely related to other measures of association like Kendall's τ, Goodman and Kruskal's γ, and Somers's d (Goodman and Kruskal, 1954; Kruskal, 1958; Somers, 1962); the differences between the measures depend on how one treats ties (Somers, 1962).

	Experiment 2		Experiment 3	
	Independent	Social influence	Independent	Social influence
Independent (exp 2)	1	0.76	0.66	0.62
Social influence (exp 2)	0.76	0.76	0.65	0.64
Independent (exp 3)	0.66	0.65	1	0.85
Social influence (exp 3)	0.62	0.62	0.85	0.90

Table 4.8: Percentage of pairs that are rank preserving. Value is concordant pairs divided by the sum of the number of concordant and discordant pairs (i.e., ties were ignored).

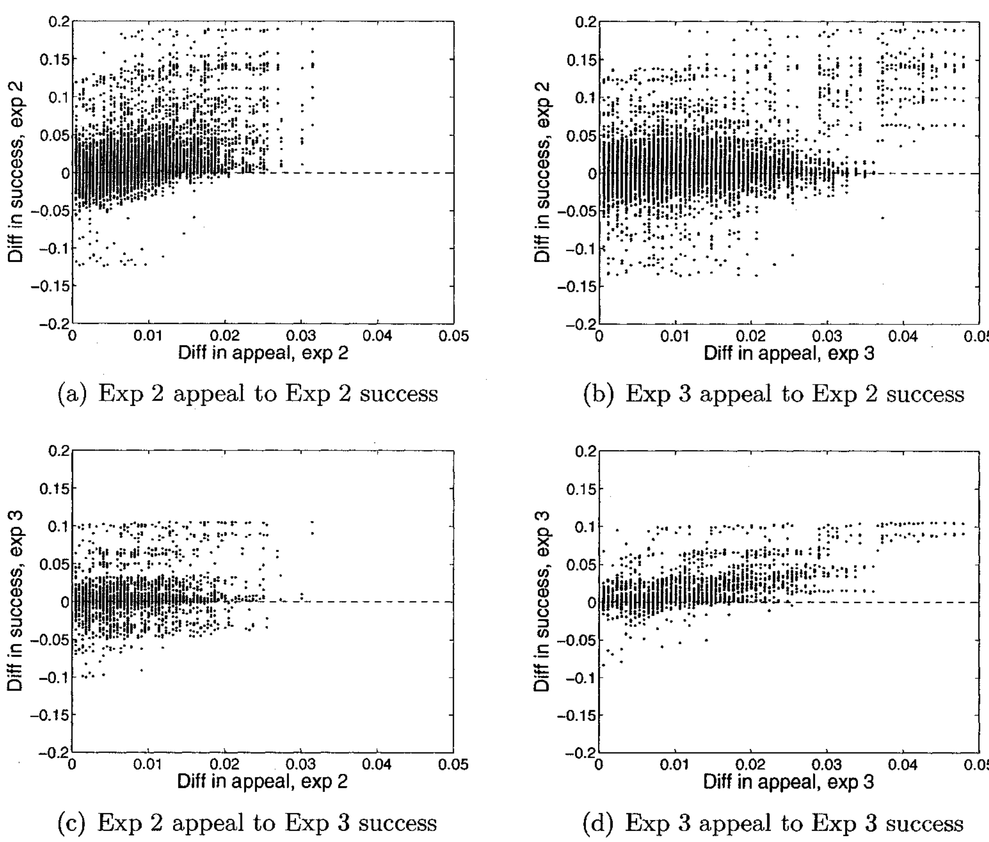
social influence world in experiment 3 than the outcomes from the independent world in experiment 3. Additionally, in experiment 2, when trying to predict an outcome in the social influence condition, it doesn't seem to matter if you have an outcome from the independent or social influence condition. The explanation for these findings in not currently known.

In addition to asking if pairs are rank preserving we can also plot the difference in success as a function of the difference in appeal (figure 4.18). In general, within experiments, as the difference in appeal increases, the expected difference in success increases. Across experiments, the pattern is still present, but more noisy. All points above the x-axis are rank preserving and points below are not.[19]

4.6 Conclusion

Returning again to the main question that motived experiment 3, we showed in section 4.3 that experiment 3 replicated the results of experiment 2 using a different pool of participants. Section 4.4 also pointed out come important differences between the two experiments which deserve further attention. We concluded the chapter by attempting in various ways to predict the success in one outcome based on information

[19]Recall that table 4.8 presents the percentage of points that are above the x-axis (not including points on the x-axis or y-axis).



(a) Exp 2 appeal to Exp 2 success

(b) Exp 3 appeal to Exp 2 success

(c) Exp 2 appeal to Exp 3 success

(d) Exp 3 appeal to Exp 3 success

Figure 4.18: Difference in success as a function of difference in appeal in experiments 2 and 3.

from another outcome. In general, we found that the independent world is useful for prediction of outcomes in a social influence world within an experiment, but not across experiments. Further, success outcomes in one experiment do not help predict success in the other experiment.

In the first three experiments, we let the success of the songs emerge naturally. In the next chapter, we turn our attention to experiment 4 where we explored the possibility of intervening in cultural markets in order to create "self-fulfilling prophecies."

Chapter 5

Experiments 3 and 4

In the first three experiments we let the success of the songs develop naturally without any intervention. In the four experiment we manipulated the market information to explore the possibility of self-fulfilling prophecies in cultural markets.

5.1 Introduction

Consumers often have information about the popularity of cultural products such as books, movies, and music from a variety of sources: friends, box office receipts, best-seller lists, and increasingly online forums. Previous work has demonstrated that information about the decisions of others tends to influence individual consumer's decisions (Hanson and Putler, 1996; Sorensen, 2007; Salganik et al., 2006; Senecal and Nantel, 2004; Huang and Chen, 2006).[1] This influence occurs in part because consumers may use the popularity of products as a signal of their quality (Hanson and Putler, 1996) and in part because consumers may benefit from coordinating their choices that is, listening to, reading, and watching the same things as others (Adler,

[1] In addition to being influenced by the gross popularity of objects, consumers are also influence by others' opinions about those objects (Chevalier and Mayzlin, 2006) and by recommendation engines (Senecal and Nantel, 2004).

1985).

The demonstrated importance of social influence in cultural markets raises the question of whether this influence can be harnessed to generate desired outcomes. Here we study whether interventions in cultural markets can generate "self-fulfilling prophecies" in which the false perception of popularity can impact consumer behavior in such a way that this initially false information becomes true (Merton, 1948, 1995).[2] Can the mere appearance of popularity, in other words, lead to actual popularity? And if so, what implications does the manipulation of information regarding the popularity of individual objects have for the market as a whole?

In spite of a few clever studies (Hanson and Putler, 1996; Sorensen, 2007), clear answers to these questions are elusive, in part because empirical evidence for self-fulfilling prophecies is generally restricted to observational data, which is invariably open to conflicting interpretations. For example, the authors Treacy and Wiersema weer accused of strategically purchasing thousands of copies of their own book, *The Discipline of Market Leaders*, in an effort to push in onto the New York Times best-seller list and thereby trigger a cascade of real sales (Stern, 1995). The book did indeed appear on the list, and once there, remained a best-seller for fifteen weeks. But how much of this success was due to the intrinsic attributes of the book itself (as the authors claim) and how much was due to a self-fulfilling prophecy? Unfortunately, answering the question would require comparing actual books sales to the sales that *would have occurred* in the absence of any manipulation, an outcomes that

[2] The concept of self-fulfilling prophecy has become clouded in the literature so we think it is useful to return to Merton's original definition: "The self-fulfilling prophecy is, in the beginning a *false* definition of the situation evoking a new behavior which makes the originally false information come *true* [emphasis in original]." Although a complete review of subsequent research in this area is beyond the scope of this footnote, it is important to point out that within psychology and education, two fields where the topic is most researched, researchers generally do not restrict the initial expectation to be false, a key difference from Merton's original definition. For example, Jussim (1986) writes, "In general, the concept of self-fulfilling prophecy refers to situations in which one person's expectations about a second person lead the second person to act in ways that confirm the first person's original expectation."

by definition cannot be observed. This problem of comparing realized with unrealized outcomes fundamentally limits the used of observational data to identify self-fulfilling prophecies (Holland, 1986).[3]

Traditional experimental methods, moreover, are also poorly suited to studying self-fulfilling prophecies. Field experiments, in which experimenters could conceivably manipulate the perceived popularity of objects in real markets, would provide the most definitive answers but such experiments would likely be expensive to conduct and many raise ethical concerns.[4] Laboratory experiments, where the manipulation of social information is possible, are well suited to studying the effects of false information on individual choice (Asch, 1952; Bond and Smith, 1996). However, in order to understand the dynamics interaction between individual behaviors and collective outcomes that are central to self-fulfilling prophecies at least thousands of subjects are required—a scale that is unfeasible in a physical laboratory.

In this chapter, therefore, we have explored another web-based experiment. Using the "multiple-worlds" design, described in detail below, we are able to observe outcomes both in the presence and absence of false information and thus measure the effect of initially false information on subsequent behavior both of individuals and the market as a whole.

[3]One way around the problems with observational data is the use of "natural" experiments. For example, Sorensen (2007) used errors in the creation of the New York Times best-seller list to attempt to get a measure of the causal effect of being on the list. While such approaches are certainly worthwhile, they are inherently limited because of the difficultly of located such natural experiments. Further, replication, extension, and elaboration of findings difficult because researchers have no control over the "design" of natural experiment.

[4]These field experiments need not be expensive when conducted online. For example, Hanson and Putler (1996) created 27 matched pairs of downloadable shareware programs and then inflated the download count of one randomly chosen program in each pair. This design yielded some clear results, but it is unlikely that the authors of the programs that were disadvantaged by this manipulation would have given consent for this experiment. One may wonder whether our bands would have given consent to have their download counts manipulated. It is the case that our consent form does not specifically mention manipulation, but, unlike the experiment of Hanson and Putler, all participating bands benefited. That is, even bands that were hurt by the manipulation were getting more attention than they would have in the absence of our experiment.

5.2 Experimental set-up

During the set-up phase of the experiment, 2,211 participants were randomly assigned to either the social influence world ($n = 752$) or the independent world ($n = 1,459$).[5]

After the popularity ordering, as measured by download counts, in the social influence world had approximately reached a steady-state, we explored the possibility of self-fulfilling prophecies by creating two new social influence worlds in which the popularity of the songs was inverted. More specifically, in both inverted worlds we set the initial download counts by inverting the ordering of the 48 songs. For example, at the time of the inversion, "She Said" by Parker Theory had the most downloads, 128, and "Florence" by Post Break Tragedy had the fewest, 9.[6] When constructing the initial conditions for both inverted worlds we swapped these songs so that participants had the false belief that "She Said" had only 9 downloads and "Florence" had 128. In this way we also swapped the 47th and 2nd songs, 46th and 3rd songs, and all other pairs of songs. The download counts for all songs immediately before and after the inversion are shown in table 5.1 and plotted for each song in figure 5.1.[7]

This manipulation most severely effected songs that were most and least pop-

[5]Approximately twice as many participants were assigned to the independent world because the set-up phase was embedded in experiment 3 (previous chapter) which was an attempt to replicate the results of experiment 2.

[6]"Florence" by Post Break Tragedy was actually tied for fewest downloads with ten other songs. This tie, and all other ties in the inversion, were broken randomly. The fact that these ties were broken randomly can actually be thought of a mini-experiment inside the bigger experiment and is consider more in section D.6.

[7]We would like to note that these manipulations are related to, but conceptually distinct from, "payola," the secret payments made to DJs in exchange for increase radio play (Randle, 1961a,b; Coase, 1979). While increased airplay of songs may lead to the misperception of popularity, it also has the potential to actually increase the amount the people like the song. Many listeners have had the experience of having an initially unfamiliar song "grow on you." Psychologists have studied this process in the large literature on familiarity and liking (Russell, 1986, 1987; Bornstein, 1989; North and Hargreaves, 1997; Cutting, 2003). Thus, "payola" involves psychological mechanisms which are not operating in this study because subjects are not repeated exposed to the same stimulus.

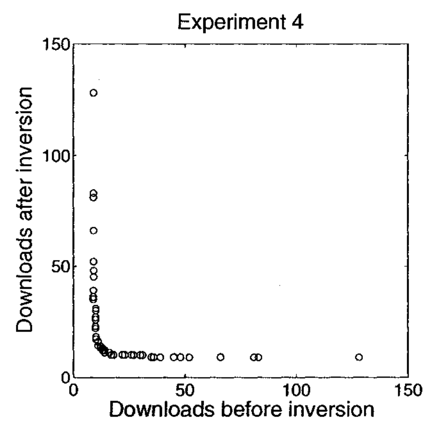


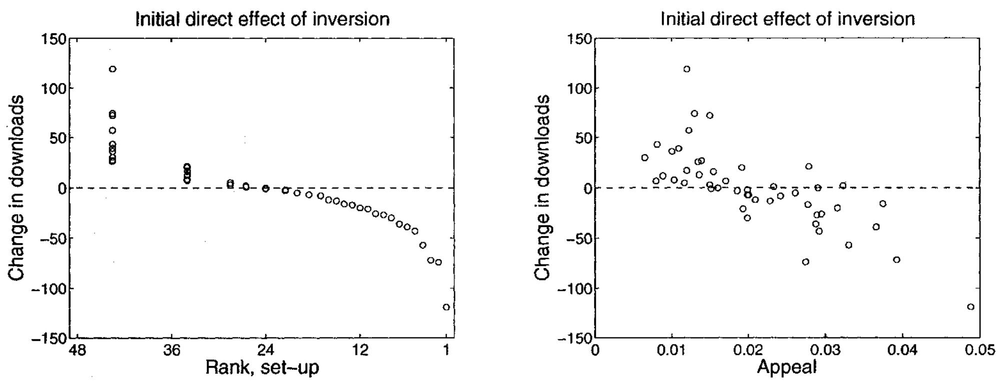
Figure 5.1: Relationship between download count before and after the inversion for the 48 songs.

ular, and because popularity was correlated with appeal, the best and worst songs (figures 5.2(a) and 5.2(b)). After this one time introduction of false information, all download counts were updated accurately. This entire process is described in figure 5.3.[8]

5.3 Subject recruitment and behavior

The set-up period 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 (Dodds et al., 2003). After the set-up period was complete, the experiment continued until August 10, 2005 (126 days) during which time we sent emails to all

[8]We are aware that a complete inversion is just one of many possible manipulations that we could have done. When choosing the manipulation we originally planned one that could be parameterized; for example, we would have had a parameter which would determine how manipulated the download counts were and then we could compare worlds with small, medium, and large distortions. However, the problem with this approach is that we needed a large number of subjects for each world so we decided that we could only have three social influence worlds. Given this number of worlds, we wanted to have one unchanged world and two copies of the manipulated world so that we could see if they both returned to the same state. Given these restrictions we could only have one manipulation and so we picked a complete inversion as the simplest manipulation which represents an extreme case. With more subjects we would certain have done others.



(a) Direct effect of inversion as a function of rank during set-up period

(b) Direct effect of inversion as a function of appeal

Figure 5.2: Direct effect of inversion as a function of rank during the set-up period and appeal. The most and least popular songs were most affected by the inversion, and because appeal and success during the set-up period were correlated, the inversion tended to most seriously affected the highest and lowest appeal songs.

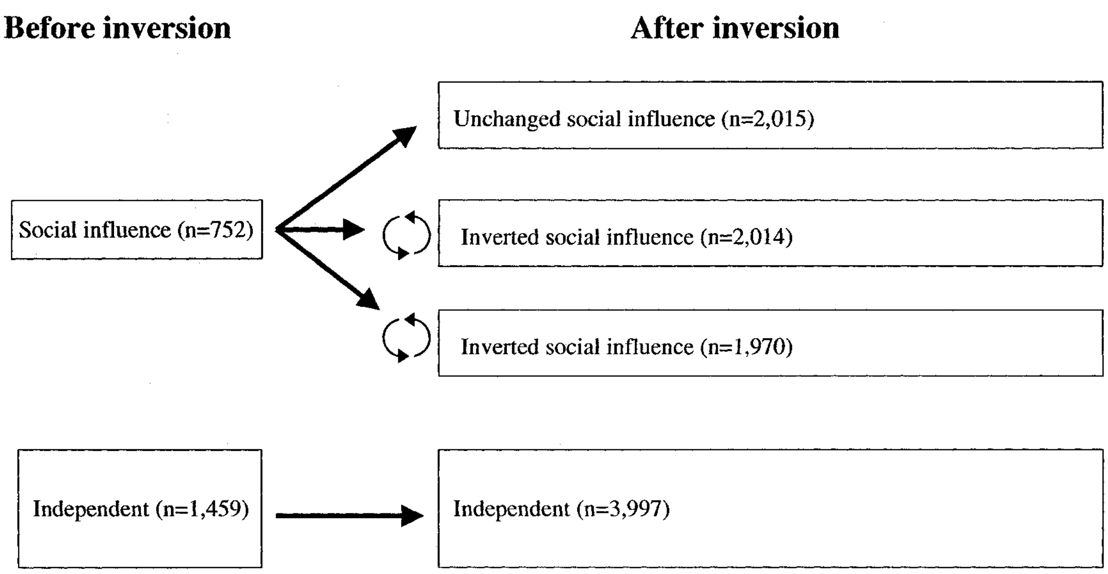


Figure 5.3: Schematic of the design of experiment 4. The area of each region is proportional to the number of subjects in that world.

Band name	Song name	Downloads before inversion	Downloads after inversion
Parker Theory	She Said	128	9
Simply Waiting	Went with the Count	83	9
Not for Scholars	As Seasons Change	81	9
Shipwreck Union	Out of the Woods	66	9
Sum Rana	The Bolshevik Boogie	52	9
Dante	Life's Mystery	48	9
Ryan Essmaker	Detour_(Be Still)	45	9
Hartsfield	Enough is Enough	39	9
By November	If I Could Take You	36	9
Star Climber	Tell Me	35	9
52metro	Lockdown	31	10
Stranger	One Drop	30	10
Forthfading	Fear	27	10
Silverfox	Gnaw	26	10
Selsius	Stars of the City	23	10
Hydraulic Sandwich	Separation Anxiety	22	10
Undo	While the World Passes	18	10
Hall of Fame	Best Mistakes	17	10
The Thrift Syndicate	2003 a Tragedy	17	10
Unknown Citizens	Falling Over	16	11
Beerbong	Father to Son	14	11
The Fastlane	Til Death do us Part (I don't)	14	12
Evan Gold	Robert Downey Jr.	13	12
Ember Sky	This Upcoming Winter	13	13
Miss October	Pink Aggression	13	13
Silent Film	All I have to Say	12	13
Stunt Monkey	Inside Out	12	14
Far from Known	Route 9	11	14
Moral Hazard	Waste of my Life	11	16
Nooner at Nine	Walk Away	10	17
Sibrian	Eye Patch	10	17
Drawn in the Sky	Tap the Ride	10	18
Art of Kanly	Seductive Intro, Melodic Breakdown	10	22
Fading Through	Wish me Luck	10	23
Benefit of a Doubt	Run Away	10	26
Salute the Dawn	I am Error	10	27
Cape Renewal	Baseball Warlock v1	10	30
Go Mordecai	It Does What Its Told	10	31
The Broken Promise	The End in Friend	9	35
Summerswasted	A Plan Behind Destruction	9	36
Secretary	Keep Your Eyes on the Ballistics	9	39
The Calefaction	Trapped in an Orange Peel	9	45
A Blinding Silence	Miseries and Miracles	9	48
Up Falls Down	A Brighter Burning Star	9	52
This New Dawn	The Belief Above the Answer	9	66
Up for Nothing	In Sight Of	9	81
Deep Enough to Die	For the Sky	9	83
Post Break Tragedy	Florence	9	128

Table 5.1: The 48 bands and songs used in experiment 4 along with the download counts before and after the inversion. During the inversion, ties were broken randomly.

remaining participants of the electronic small-world experiment who had not been contacted during the set-up period ($n = 50,800$). Because the emails led to web-postings about the experiment, there were also participants who were not associated with the electronic small-world experiment.

Table 5.2 presents demographics of the participants. Generally, the demographics were similar before and after the inversion with the exception of the large increase in the number of Brazilians which was caused by a mention of the experiment on the popular Brazilian website `http://www.estadio.com.br`.

Tables 5.3 and 5.4 present simple descriptive statistics of the subjects' behavior. Participant behavior before and after the inversion was also similar, with the exception of the listening and download behavior in the inverted worlds, as will be described more fully in section 5.5. Figures 5.4 and 5.5 plot the distribution of listens and downloads per participant in experiment 4.

The experiment 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.

5.4 Was the experiment stable?

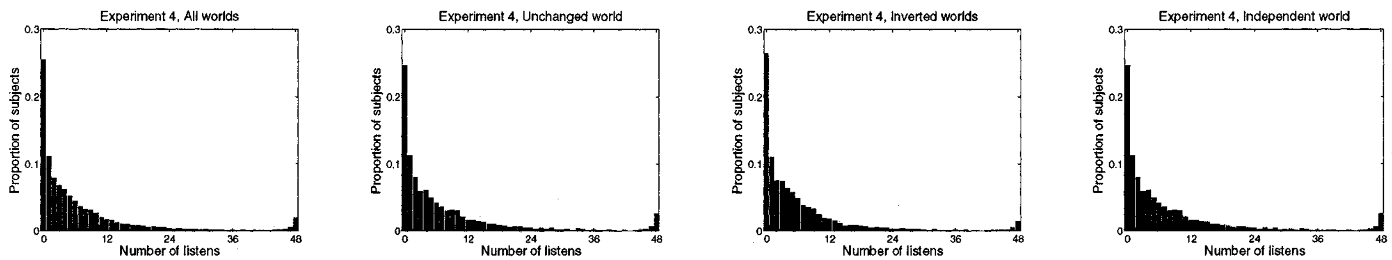
Because the experiment ended for exogenous reasons, it is important to ask if the worlds were actually stable at the time the experiment ended. A natural first check is to plot the download counts for all songs over time in all three worlds. As shown in figure 5.6, these download counts are always increasing so it is somewhat difficult to eyeball a measure of stability. An easier way to detect stability would be to examine the market share values over time as seen in figure 5.7. To the eye the

Category	Before inversion ($n = 2,211$) (% of participants)	After inversion ($n = 9,996$) (% of participants)
Female	37.9	43.9
Broadband connection	91.3	89.4
Has downloaded music from other sites	69.4	65.3
Country of Residence		
United States	68.1	54.7
Brazil	1.2	12.5
Canada	6.8	4.9
United Kingdom	6.6	6.9
Other	17.3	21.0
Age		
14 and younger	1.4	2.3
15 to 17	5.4	5.6
18 to 24	29.6	26.6
25 to 34	35.8	33.9
35 and older	27.8	31.7

Table 5.2: Descriptive statistics about the participants before and after the inversion. Totals many not sum to 100 due to rounding.

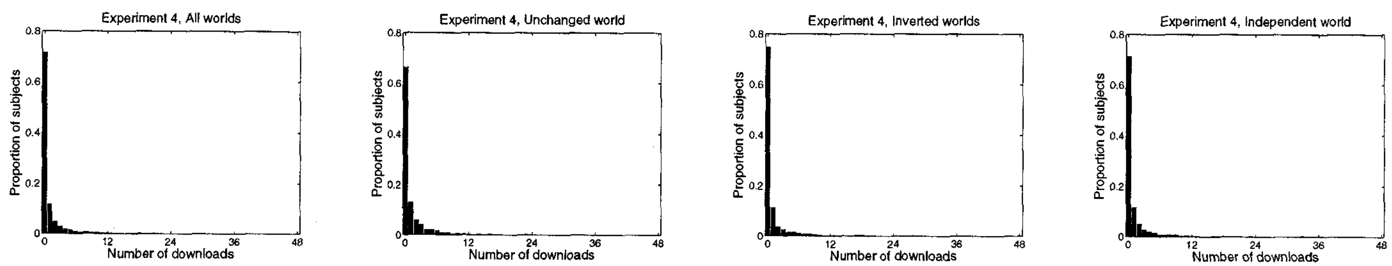
	Before inversion		
	Social influence ($n = 752$)	Independent ($n = 1,459$)	Total ($n = 2,211$)
Number of listens	5,628	11,844	17,472
Mean per subject	7.5	8.1	7.9
Number of downloads	1,133	1,691	2,824
Mean per subject	1.5	1.2	1.3
Pr[download | listen]	0.201	0.143	0.162
Average rating (# of stars)	2.70	2.55	2.62

Table 5.3: Descriptive statistics of subject behavior before the inversion.



(a) Exp 4, all worlds (b) Exp 4, unchanged world (c) Exp 4, inverted worlds (d) Exp 4, independent world

Figure 5.4: Distribution of the number of listens per participant in experiment 4. Results are also presented for just participants in the unchanged world, inverted worlds, and independent world.

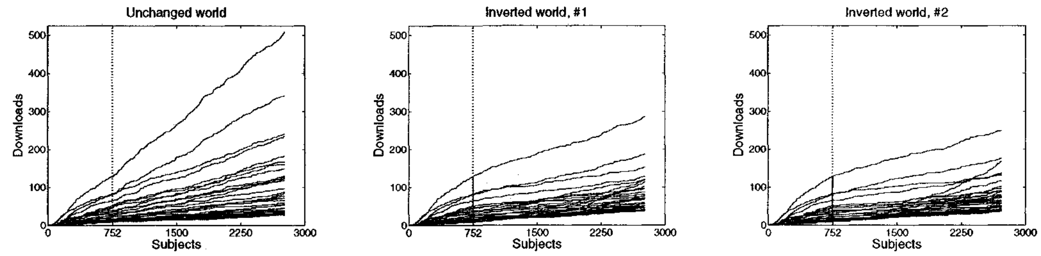


(a) Exp 4, all worlds (b) Exp 4, unchanged world (c) Exp 4, inverted worlds (d) Exp 4, independent world

Figure 5.5: Distribution of the number of downloads per participant in experiment 4. Results are also presented for just participants in the unchanged world, inverted worlds, and independent world.

	After inversion				
	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 5.4: Descriptive statistics of subject behavior after the inversion. These download counts do not include the initial conditions.



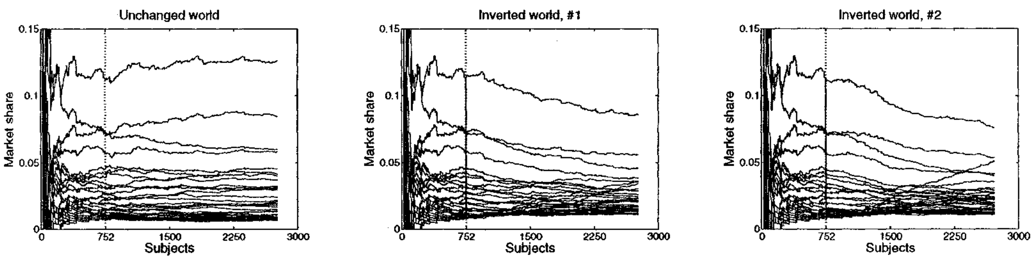
(a) Unchanged world (b) Inverted world (c) Inverted world

Figure 5.6: Download counts for all songs over time before and after the inversion. Some of the songs are clearly not stabilized especially in the second inverted world.

market share of the songs in the unchanged world look pretty stable, but there are some songs in the second inverted world that are clearly not stable. These download and market share dynamics can also be compared to the dynamics in the independent world (figure 5.8).

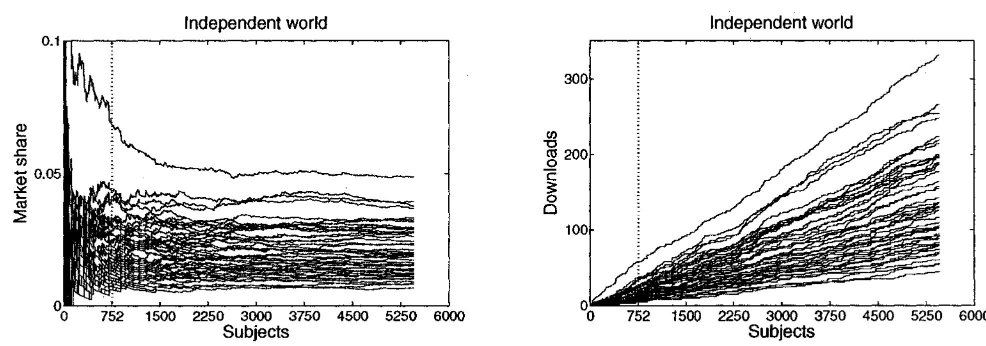
Rather than just eyeballing the market share trajectories in figure 5.7, we can also plot the slope market share trajectories over the last x subjects, where we will call x the window size, here 1,000 (other window sizes produce similar results). If these slopes are 0 (or close to it) then the market shares are stable. To estimate these slopes we made a linear-least squares fit to the trajectory over the last x subjects.[9] Since these slopes are very small numerically, we multiplied them by 1,000 to yield

[9]Note that this assumes that the trajectories are linear and not curvilinear. Over the last 1,000 subjects most of the market share trajectories are reasonably linear (see figure 5.7).



(a) Unchanged world (b) Inverted world (c) Inverted world

Figure 5.7: Market share for all songs over time before and after the inversion. Some of the songs are clearly not stabilized especially in the second inverted world.



(a) Independent world – market share (b) Independent world – downloads

Figure 5.8: Download count and market share in the independent condition for all songs over time before and after the inversion.

the projected change in market share over the next 1,000 subjects.[10] Figure 5.9 plots the slope of the market share as a function of rank at the end of experiment 3 in all four worlds. Recall that if the market share for a particular song was stable then this slope would be 0. In the unchanged and independent worlds the slopes, indeed, seem very close to 0 for almost all songs. However, in the inverted worlds, the slopes for the original top songs seem to be positive indicating that they are gaining market share (as can also be seen in figures 5.7(b) and 5.7(c)). Thus, it seems that at least some songs in the inverted worlds had not yet stabilized. Interpreting the numerical magnitude of these values is difficult. In the plots the y-scale is $2 \cdot \frac{1}{48}$ or two times the market share of the average song. Even the songs that are changing most quickly seem to change about $\frac{1}{48}$ per 1,000 subjects.

5.4.1 Projection-based measures of stability

Yet another way to explore the stability is to project into the future. This projected future outcome can then be compared with the final outcome that we did observe. To make these projections we followed a similar logic to procedure looking at stability and calculated a linear least-squares fit to the download trajectories over a given window size.[11] For example, figure 5.10 plots the download trajectory for the song "She Said" by Parker Theory in inverted world #1 along with the least-squares line for two different window sizes: 500 and 1,000. These different window sizes yield slightly different slopes: 79 and 84 downloads per 1,000 subjects.[12] Note that this

[10] It is just a coincidence that the window size and the rescaling factor are both 1,000. For the rescaling factor we picked 1,000 because it produced slopes in the range of 0 to 0.04 which are easier to comprehend because they can be compared to the market share of the average song which is $\frac{1}{48} \approx 0.02$.

[11] There is no particular reason that we made the projections based on the download trajectories rather than the market share trajectories. Either would have worked.

[12] Again, it is just a coincidence that the rescaling factor is 1,000. We picked this value because it produces download counts in the range of 0 to 100.

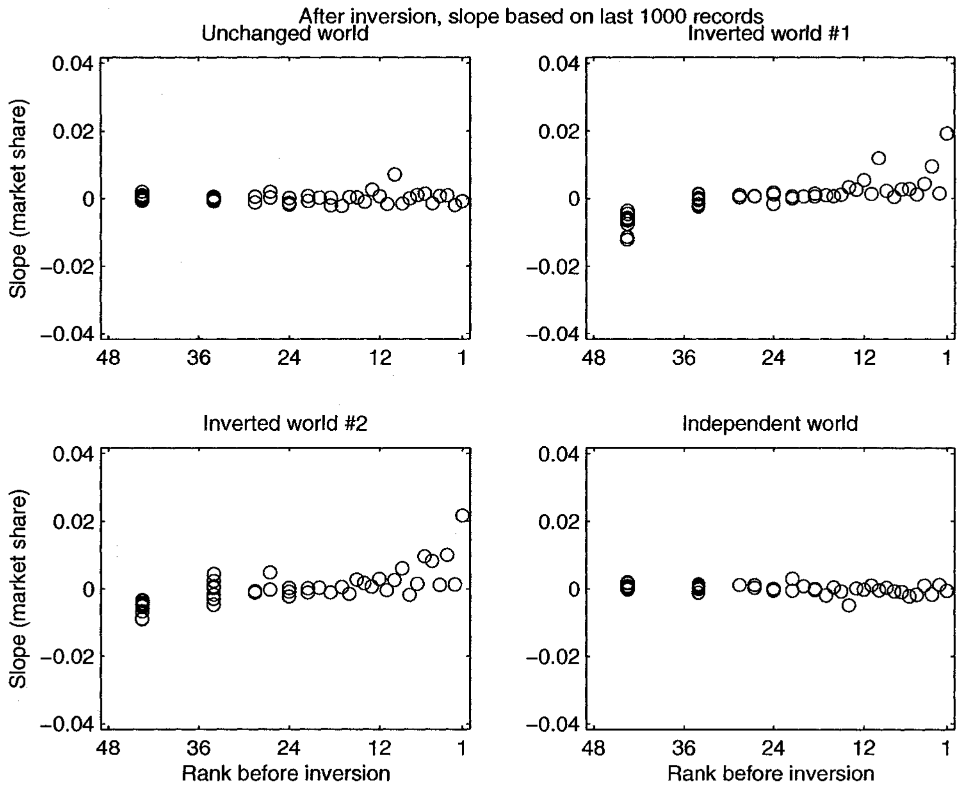


Figure 5.9: Slope of market share trajectories in the four worlds in experiment 4 after the inversion. The upper left figure is the unchanged world. The upper right and lower left figure are the inverted worlds. The lower right figure is the independent condition. The market shares seems to be pretty stable in the unchanged and independent worlds, but the not the inverted worlds.

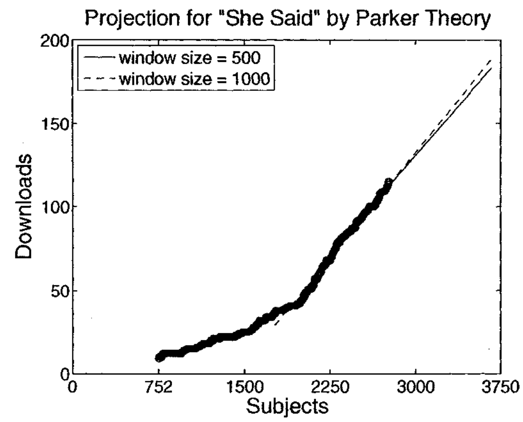


Figure 5.10: Projected download growth for "She Said" by Parker Theory in inverted world #1. The two different lines are based on fits with different window sizes.

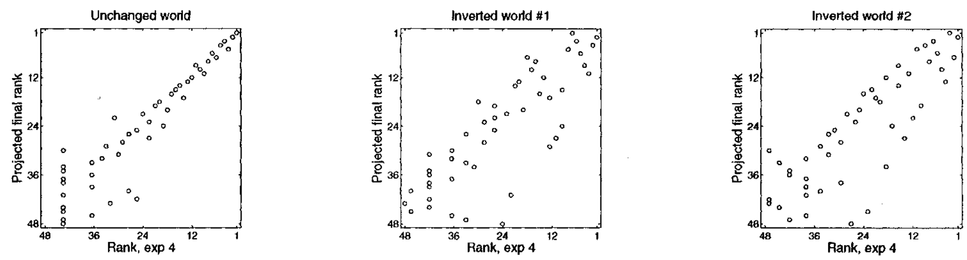
projection procedure assume that the download trajectories will continue to change in a linear way.

To project the long-run state of the system we can compare the slopes for the songs because over long timescales this rate of growth will overwhelm any difference introduced by the initial conditions that we created. Figure 5.11 plots the actual rank at the end of experiment 4 compared to the projected rank. If the current state and projected final state were the same all points would fall along the 45-degree line. This is approximately what we see in the unchanged world (figure 5.11(a)), but not the inverted worlds (figures 5.11(b) and 5.11(c)).

To conclude, the unchanged world seemed stable, but the inverted worlds were still changing some, particularly the songs that were highest rank before the inversion. Now we turn our attention to the results.

5.5 Results: Individual behavior

During the experiment participants listened to about 7 songs, on average, and downloaded 1; tables 5.3 and 5.4 report simple descriptive statistics about subject



(a) Unchanged world (b) Inverted world (c) Inverted world

Figure 5.11: Comparing final rank and projected final rank in experiment 4. In the unchanged world, the rank ordering is similar $(r = 0.92)$, however the inverted worlds the songs have not completely moved to their new ordering: $(r = 0.83)$ and $(r = 0.84)$.

behavior. One interesting feature of these data is that subjects downloaded substantially fewer songs in the inverted worlds than the unchanged world: 2,197 and 2,160 compared with 2,898. The total number of downloads is based on two components: the total number of listens and the probability of download given listen:

$$Pr[\texttt{download}] = Pr[\texttt{listen}] \cdot Pr[\texttt{download} \mid \texttt{listen}] \tag{5.1}$$

In the inverted worlds both of these factors were lower than in the unchanged world. Next we will explore the reasons for these differences.

In this experiment, subjects' decisions about which song to listen to where affected the the perceived popularity of the songs. At the time each subject participated, every song in their world had a specific download count and market rank (for example, the song with the most downloads had a market rank of 1). To measure the effect of the popularity on the listening decision, we can calculate the probability that a subject listened to the song of a given market rank (independent of which song occupied that rank at the time). Figure 5.12 shows that in both the unchanged world and the inverted worlds, subjects were more likely to listen to whichever song they believed was more popular with a reversal of this pattern at the very bottom of the popularity ranking—i.e., subjects were more likely to listen to then 48th song rather

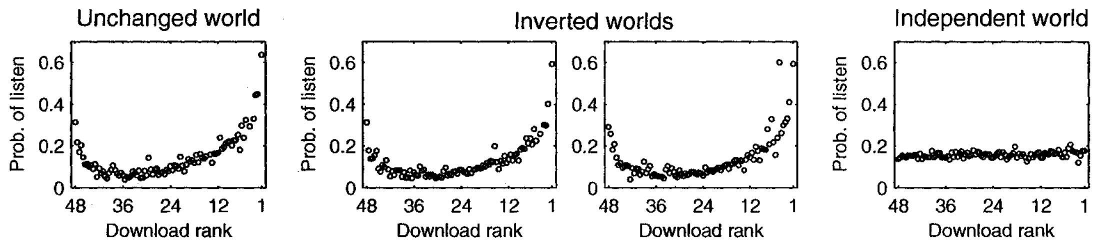


Figure 5.12: Probability that a subject in each world listened to a song of a given market rank in experiment 4. These plots are not smoothed.

than say the 45th.[13] For example, compared to the song of middle popularity (rank 24 out of 48), participants in the influence worlds were about six-times more likely to listen to the most popular song and three-times more likely to listen to the least popular. The behavior of the subjects in the three social influence worlds can be sharply contrasted with the behavior in the independent world where the probability of listening to a song was not affected by its popularity (which is as it should be because this information was not available to participants).

One consequence of this listening pattern is that in the inverted worlds, the lower appeal songs tended to get more listens as can been in figure 5.13 which plots the market share of listens as a function of appeal. In the unchanged world the top 10 best songs had about twice as many listens as the 10 worst songs; in the inverted worlds this pattern was reversed.[14] Because participants in the inverted worlds were disproportionately exposed to lower appeals songs, on average, they listened to fewer songs before leaving the experiment (6.2 and 6.4 compared to 7.2) and were less likely

[13]This tendency to prefer to listen to the least popular songs was also observed in experiments 2 and 3 (see chapters 3 and 4), but the source of this pattern is unclear. It could be an artifact of the list format used in experiment which, in addition to making the top-ranked songs more salient, also made the bottom-ranked songs more salient. Or, instead of being an artifact of our experiment, this could be a real behavioral tendency for some people to want to listen to the least popular songs, perhaps as a form of anti-conformist behavior (Simmel, 1957; Heath et al., 2006). Further experiments would be required to adjudicate between these possibilities.

[14]The market share of listens for the 10 best songs in the three worlds were 0.39, 0.17, and 0.17. The market share of listens for the 10 worst songs were 0.14, 0.30, 0.30.

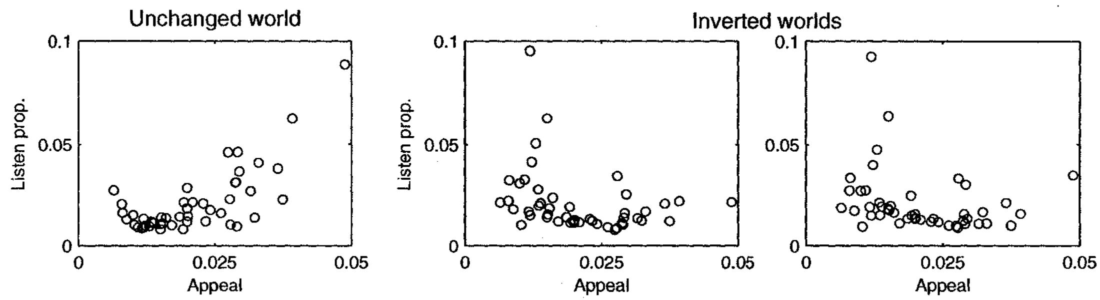


Figure 5.13: Relationship between appeal and number of listens in experiment 4. In the inverted worlds the lower appeal songs get more listens.

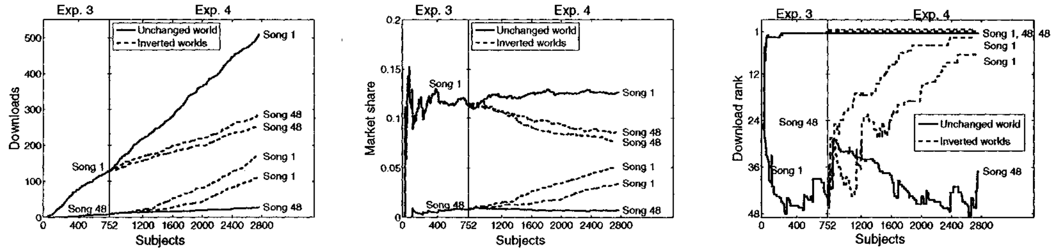
to download the songs to which they listened (0.176 and 0.171 compared to 0.201).[15] Together these two changes lead to the reduction in downloads.

5.6 Results: Pairwise comparison of songs

A natural first analysis of system dynamics would be to compare the songs that were swapped. Thus, we could compare the songs that was 1st and the song that was 48th. However, as noted previously, and as shown in table 5.1, there were ties so the 48th song is not uniquely defined. We will call the song that was swapped into 1st place the 48th song.

Figure 5.14 plots the download, market share, and market rank trajectory of song 1 ("She Said" by Parker Theory) and song 48 ("Florence" by Post Break Tragedy). In figure 5.14(a) we see that the download trajectories in the unchanged world (solid line) were similar before and after the inversion, but the trajectories in the inverted worlds (dashed worlds) were quite different. Song 48 earned downloads at a faster rate as a consequence of the inversion (i.e. the slope of the download trajectory was steeper) and song 1 earned downloads at a slower rate. Thus both songs seemed

[15]Another way to measure the increased dissatisfaction that was caused by the inversion is to note that the average rating (# of stars) in the unchanged world, 2.71, was greater than in either inverted world, 2.64 and 2.60.

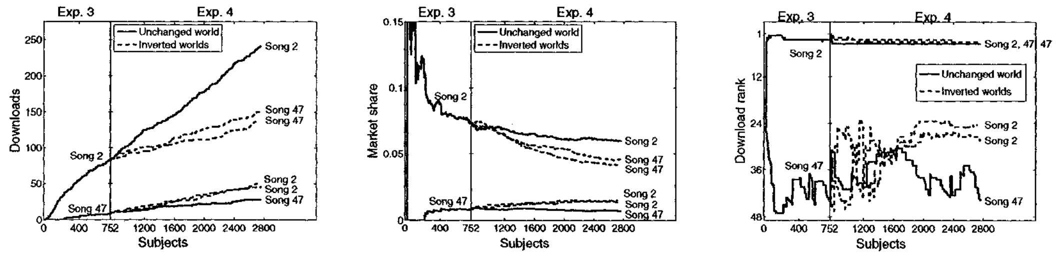


(a) Downloads (b) Market share (c) Download rank

Figure 5.14: Success dynamics of song 1 ("She Said" by Parker Theory) and song 48 ("Florence" by Post Break Tragedy).

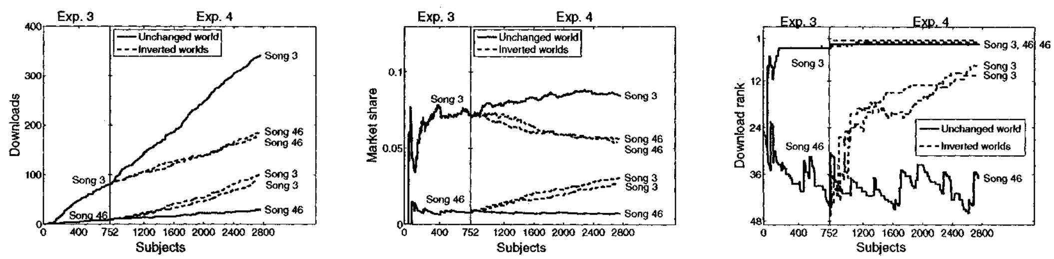
to have experienced a self-fulfilling prophecy where the perception of success affected future success. On account of the finite number of subjects in experiment, however, it is impossible to say with certainty whether these self-fulfilling dynamics would have had a permanent effect on the popularity of the songs. Indeed, the relative slopes of songs 48 and 1 in the inverted worlds suggest that song 1 may eventually catch-up to song 48 (more on that in section 5.6.1). Figure 5.14(b) plots the market share trajectories for these two songs. The market share of both songs in the unchanged world is similar to the values before the inversion, but the results in the inverted world are different. Again, it looks as if song 1 will overtake song 48. Figure 5.14(c) plots the rank trajectories which show that in both the unchanged and inverted world, the song that started experiment 4 ranked first, stayed first during the entire experiment (again probably because of the finite number of subjects). However, we can see that in the inverted world song 1 rapidly gained rank.

Figure 5.15 plots the download, market share, and market rank trajectory for song 2 ("Went with the Count" by Simply Waiting) and song 47 ("For the Sky" by Deep Enough to Die). These results are generally similar to those from figure 5.14 except it appears to be the case that song 2 will not catch up to song 47. Thus, for this pair of songs the inversion seems completely self-sustaining. Figure 5.16 plots the download, market share, and market rank trajectory for song 3 ("As Seasons



(a) Downloads (b) Market share (c) Download rank

Figure 5.15: Success dynamics of song 2 ("Went with the Count" by Simply Waiting) and song 47 ("For the Sky" by Deep Enough to Die).



(a) Downloads (b) Market share (c) Download rank

Figure 5.16: Success dynamics of song 3 ("As Seasons Change" by Not for Scholars) and song 46 ("In Sight Of" by Up for Nothing).

Change" by Not for Scholars) and song 46 ("In Sight Of" by Up for Nothing) which again are broadly similar.

5.6.1 Robustness of pairwise comparison

To ascertain the projected final ordering of the pair of songs, and the robustness of this projection, we can compare their projected rate of download growth (i.e. the slope of their download trajectory)—over long timescales this rate of growth will overwhelm any difference introduced by the initial conditions that we created. Figure 5.17 plots the rate of growth for song 1 and song 48 as a function of the window size used to make the projection. In the unchanged world, song 1 clearly has a much larger rate of growth than song 48, independent of window size. In one of

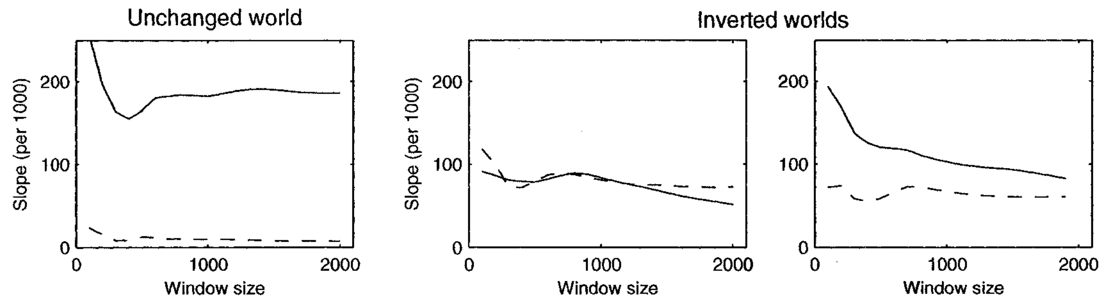


Figure 5.17: Projected slope per thousand subjects for song 1 ("She Said" by Parker Theory) (solid line) and song 48 ("Florence" by Post Break Tragedy) (dashed line) as a function of window size. In the unchanged world, song 1 is projected to have many more downloads than song 48, independent of window size. In one of the inverted worlds, the projected order of these two songs depends on the window size used, but in the other inverted world, song 1 is projected to finish ahead of song 48 for all window sizes.

the inverted worlds, the specific window size affects which song has the greater slope, and thus affects our substantive conclusion about the final ordering of the songs. In the other inverted world, however, song 1 has a greater slope for all window sizes. For songs 2 and 47 the results are robust to the window size used for the projection (figure 5.18). That is, no matter which window size is used, in the unchanged world song 2 is always projected to be ahead of song 47 and in both inverted worlds song 47 is always projected to be ahead of song 2.

In a similar manner we checked the relatives slopes in the inverted worlds for the top 10 pairs of songs and the results are summarized in table 5.5. It seems that, in general, for most pairs the songs the inversion was not self-sustaining. However, this does not necessarily mean that the songs returned to their original ranking. Rather, it just means that this particular pair of songs returned to their original relative ordering.

Thus, it seems that some pairs of songs return to their original ordering and that others don't. Results exploring the timescale of these recovery (i.e. the amount of time it takes for the songs to return to their original order) are presented in section D.4

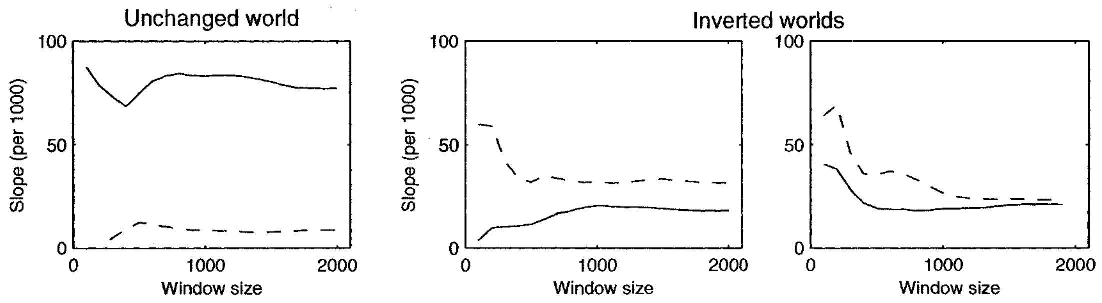


Figure 5.18: Projected slope per thousand subjects for songs 2 ("Went with the Count" by Simply Waiting) (solid line) and song 47 ("For the Sky" by Deep Enough to Die) (dashed line) as a function of window size. In the unchanged world, song 2 is projected to have many more downloads than song 47, independent of window size. However, in both inverted worlds, song 47 is projected to finish ahead of song 2 for all window sizes.

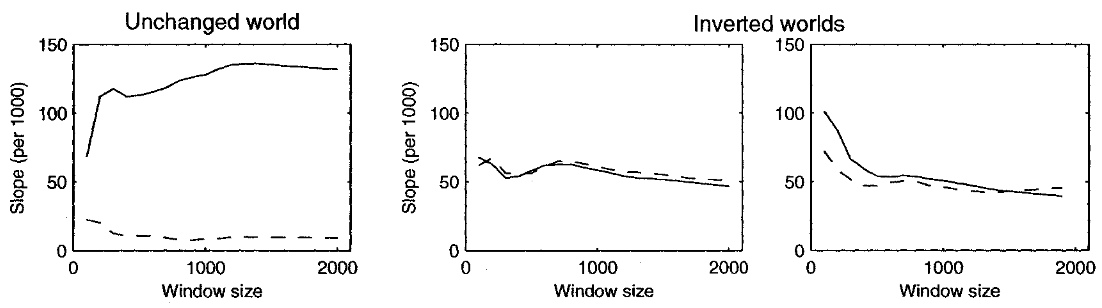


Figure 5.19: Projected slope per thousand subjects for songs 3 ("As Seasons Change" by Not for Scholars) (solid line) and song 46 ("In Sight Of" by Up for Nothing) (dashed line) as a function of window size. In the unchanged world, song 3 is projected to have many more downloads than song 46, independent of window size. However, in both inverted worlds the projected order of finish depends on the window size used for projections.

Was the inversion self-sustaining?		
Song pair	Inverted world #1	Inverted world #2
1/48	?	no
2/47	yes	yes
3/46	?	?
4/45	?	?
5/44	yes	no
6/43	no	no
7/42	no	?
8/41	no	no
9/40	no	no
10/39	no	no

Table 5.5: For most pairs of songs the inversion was not self-sustaining. The value "?" means that the results depend on the window size used for the projections.

and are generally found to be uninformative. Next we turn our attention to the dynamics of the entire set of songs.

5.7 Results: All songs

Rather than examining the success of the most and least popular songs, we can also consider the outcomes for the entire set of songs. First we consider the inequality of success, as measured by the Gini coefficient. Figure 5.20 shows that after the inversion the Gini coefficient in the unchanged world (solid line) remained approximately stable, but that in the inverted worlds (dashed lines) the Gini coefficient decreased and then appeared to stabilize. The Gini coefficient in the inverted worlds decreased because, relative to the unchanged world, the top songs earned downloads more slowly and the bottom songs are earned downloads more quickly leading to a decrease in inequality (also, see figure 5.7). The finding that Gini coefficient is approximately stable in the unchanged world is important for the findings in the previous chapters; it shows that had we run the experiment for 2,000 subjects per world instead of 700

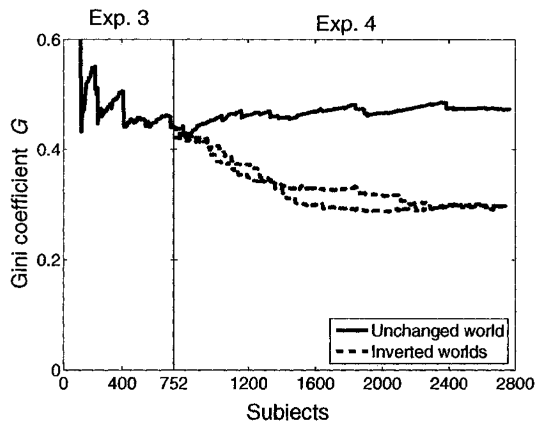


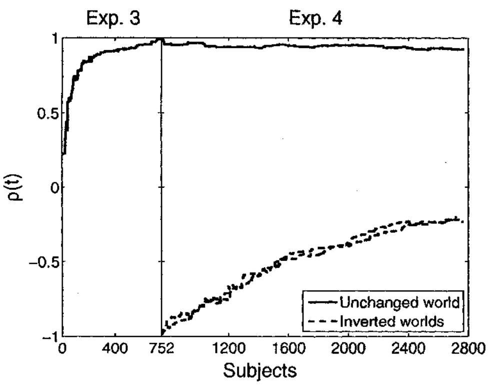
Figure 5.20: Dynamics of Gini coefficient in experiment 3 and 4. In the unchanged world the Gini coefficient was approximately stable, but in the inverted worlds it initially decreased and then seemed to stabilize.

we would have likely seen similar levels of inequality.

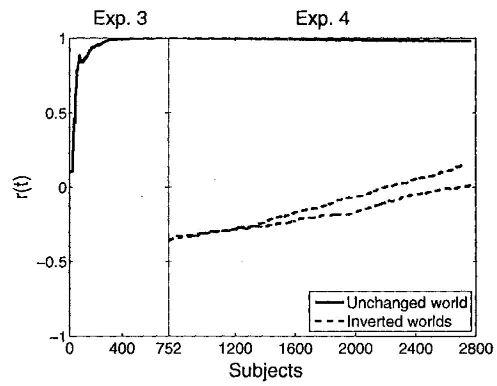
It is also useful to explore system dynamics by comparing the success after the inversion to the success during the set-up period and to the appeal of the songs. We will consider these comparisons in the next two sections.

5.7.1 Results in terms of success during the set-up period

It is natural to compare the results with the inversion with the success of the songs during the set-up period. For example, we can ask if after the inversion the songs returned to their old ordering? Figure 5.21(a) plots Spearman's rank correlation between popularity at the end of the the set-up period and and the popularity in the three social influence worlds as a function of time $\rho(t)$. Before the inversion (to the left of the vertical line) the popularity ordering quickly moved to its final state. Then after the inversion (to the right of the vertical line) the popularity ordering did not change much in the unchanged world as is indicated by the rank correlation remaining close to one. These dynamics indicate that the inversion occurred after the rankings had stabilized and serve as a control condition which can be compared the to the very



(a) Rank correlation $\rho(t)$

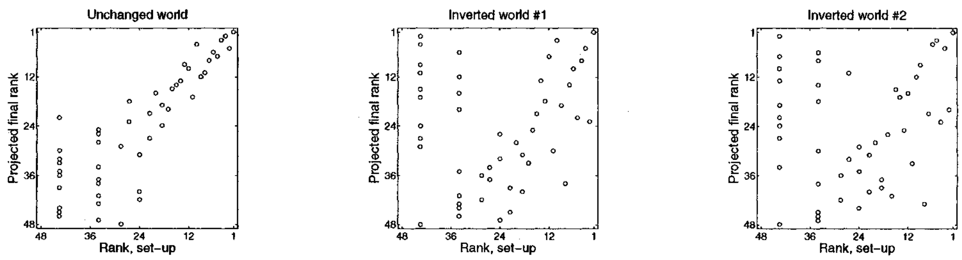


(b) Correlation $r(t)$

Figure 5.21: Spearman's rank correlation profile, $\rho(t)$, and correlation profile, $r(t)$, between popularity at the end of the set-up period and popularity in the unchanged world (solid line) and inverted worlds (dashed line). In the inverted world $\rho(t) = -1$ by definition, but $r(t) \neq -1$. However the system did not lock-in to this inverted state.

different system behavior that was observed in the inverted worlds (dashed lines). By definition the inversion caused $\rho(t) = -1$. However, neither inverted world locked-in to this new state. Instead, the songs in both inverted worlds gradually returned to their pre-inversion ordering. Thus, participants were able to "recover" at least some of the original ordering. However, it appears that this recovery would have stopped well short of returning to the original ordering.

Figure 5.21(b) plots the same results for the correlation profile $r(t)$. Qualitatively the results are similar, but there are minor differences. The correlation more quickly approaches 1 during the set-up period probably because it is less affected by the noise at the bottom of the popularity distribution. Also, the correlation does not become -1 at the time of inversion because of the nature of the inversion; see figure 5.1. In general, these two plots tell a similar story: the unchanged world seems to stay pretty stable, but the inverted worlds gradually return to their previous ordering. However, visual inspection of both figures suggests that this recovery may fall well short of returning to the original ordering.



(a) Unchanged world (b) Inverted world #1 (c) Inverted world #2

Figure 5.22: Comparing the rank at the end of the set-up period to the projected final rank. In the unchanged world, the rank ordering is similar $(\rho = 0.84)$, but the inverted worlds the songs have not recovered their original rank ordering $(\rho = 0.16, 0.16)$.

To learn more about the projected final state of the system we compare the rank during the set-up period to the projected final rank. In figure 5.22(a) we see that the initial and final ranks are highly correlated $(\rho = 0.84)$, with some noise at the bottom of the popularity distribution where a few downloads can result in a large change in rank. Figures 5.22(b) and 5.22(c) show that in the inverted worlds, the projected final rank is only weakly related to the rank before the inversion $(\rho = 0.16, 0.16)$. Thus, in these two worlds the system moves to a new state which has almost no relationship to its old state.[16] As seen in figure 5.23, these results are robust to the particular window size used for the projections.

Using these projected final ranks in each world we can calculate the benefit (or cost) of the inversion for each song by comparing its projected final rank in the inverted and unchanged worlds,

$$\Delta K_i = -(K_{i,inv} - K_{i,unc}) \tag{5.2}$$

where $K_{i,inv}$ is the projected final rank for song i in the inverted world and $K_{i,unc}$ is the projected final rank in the unchanged world.[17] A positive value of ΔK_i means

[16]Figure 5.22 may also raise questions about the relationship of the inverted worlds to each other. As shown in section D.5, the inverted worlds move into similar steady states.

[17]The minus in equation 5.2 is because a rank of 5 in the inverted world and 1 in the unchanged

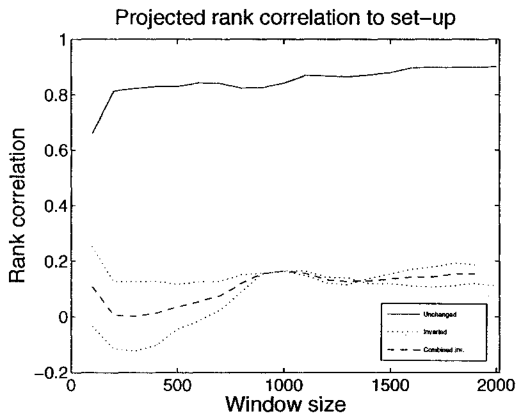


Figure 5.23: Rank correlation between set-up and projected final rank as a function of window size. The projected final state in both inverted worlds has almost no relationship to the ordering before the inversion, a results that is robust tot he window size used to make the projection.

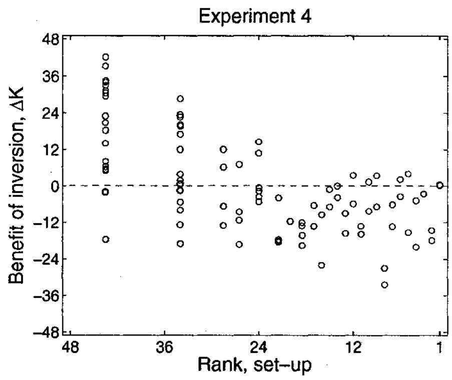
that the song was projected to have a higher rank in the inverted world. Figure 5.24(a) plots ΔK as a function of the rank during the set-up period and shows that, in general, songs that were promoted by the inversion (ranked between 25th and 48th during the set-up period) tended to have better long-term results as a consequence of the inversion. Conversely, songs that were hurt by the inversion (ranked between 1st and 24th during the set-up period), tended to have worse long-term outcomes in the inverted worlds. Thus, the long-term success of almost all songs was influenced by our one-time manipulation. However, song 1 ("She Said" by Parker Theory) is projected to return to the top spot.

Looking at the effect of the inversion simply in terms of projected final rank loses some information. We can also calculate the benefit of inversion, by comparing the slope of the download trajectories for a song in the inverted and unchanged world,

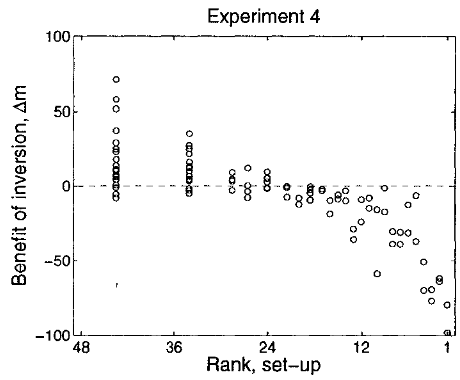
$$\Delta m_i = m_{i,inv} - m_{i,unc} \tag{5.3}$$

where $m_{i,inv}$ is the slope of song i in the inverted world and $m_{i,unc}$ is the slope in

world is actually a decrease in success.



(a) Benefit of inversion ΔK



(b) Benefit of inversion Δm

Figure 5.24: Benefit of the inversion as a function of the rank during the set-up period. There are two values for each song (one from each inverted world). In terms of rank, ΔK , and slope, Δm , songs that were promoted by the inversion tended to better in the long-run and songs that were hurt by the inversion tended to do worse. Results are jittered to resolve ties.

the unchanged world. A positive value of Δm_i means that a song earned downloads faster in the inverted world. Qualitatively, figure 5.24(b), which plots the effect of the inversion on slopes, matches the results of figure 5.24(a), which plots the effects of the inversion on ranks. They both show that songs that were promoted by the inversion tended to have better long-term results. However, looking at the results in terms of slopes (figure 5.24(b)) also reveals that even thought song 1 ("She Said" by Parker Theory) is projected not to have suffered in terms of projected final rank, it definitely suffers a lower download rate as a result of the inversion.[18] The fact that song 1 is projected to return to the top spot even thought it was severely penalized in terms of download rate naturally raise the question of whether this song was somehow special. Thus, in the next section we will consider the results in terms of appeal.

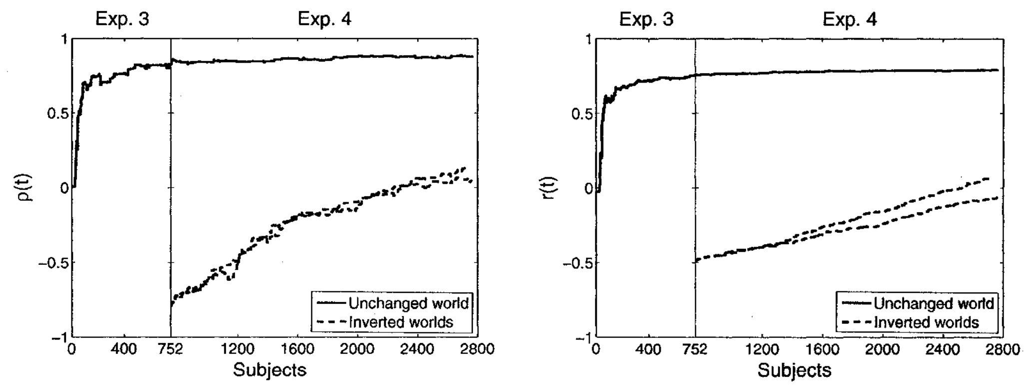
[18]This pattern of song 1, 2, and 3 having slower rates of download in the inverted worlds can also be seen in figures 5.14(a), 5.15(a), and 5.16(a).

5.7.2 Results in terms of appeal

In addition to considering the results in terms of the rank at the end of experiment 3, it is also insightful to consider the results in terms of appeal (recall that appeal is measure by the market share in the independent condition). Figures 5.25(a) and 5.25(b) mirror figures 5.21(a) and 5.21(b) by plotting the rank correlation profile, $\rho(t)$, and correlation profile, $r(t)$, for the relationship between appeal and success. Both figures show that before the inversion (to the left of the vertical line) the results quickly settle down. After the inversion (to the right of the vertical line), the unchanged world (solid line) continues its trajectory and the inverted worlds start to recover to the previous state but then seem to level-off.[19] The results in the unchanged world are important for the results from the previous chapters; it shows that the imperfect relationship between appeal and success was not caused by only having 700 subject. Rather, the results in the unchanged world suggest that even if the previous experiments had run for 2,000 subjects in each world we will still see an imperfect relationship between quality and success.

We can also use our measure of appeal to show closely the projected outcomes in the social influence worlds reflected the true preferences of our population. The projected outcomes in the unchanged world were strongly, but not completely, related to to appeal ($\rho = 0.82$) (figure 5.26(a)); thus the presence of social information produced outcomes that did not perfectly reflect the true preferences of the population. The inverted worlds, however, were much less reflective of population

[19]The fact that the highest appeal songs gradually gain popularity in the inverted world allows us a chance to return to the argument about the reduction in downloads in the inverted worlds. Recall that, as described in section 5.5, subjects in the inverted worlds downloaded fewer songs than those in the unchanged world. We argued that this was because subjects in the inverted world were disproportionately exposed to bad songs. However, as the best songs gained popularity the overexposure to bad songs could have decreased suggesting that the average number of downloads per subject in the inverted worlds could increase over time. In section D.3 we consider this possibility in detail. To summarize these findings, there is almost no support that subjects downloaded more songs as the better songs became more popular. However, additional experiments would be needed to explore this question directly.

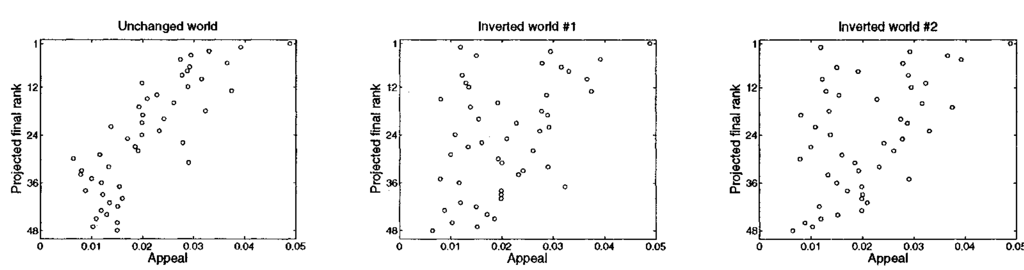


(a) Rank correlation, appeal and success $\rho(t)$ (b) Correlation, appeal and success, $r(t)$

Figure 5.25: Spearman's rank correlation profile, $\rho(t)$, and correlation profile, $r(t)$, between appeal and popularity in the unchanged world (solid line) and inverted worlds (dashed line). Neither inverted world locked-in to the inverted state.

preferences ($\rho = 0.40, 0.45$) (figures 5.26(b) and 5.26(c)), suggesting—perhaps not surprisingly—that markets in which perceived popularity has been manipulated will in general be less revealing of true preferences than markets in which popularity is allowed to emerge naturally. These results are robust to the window size used for projecting the final ranks (figure 5.27).

Finally, we can calculate the benefit of inversion as a function appeal. Figure 5.28 plots the benefit, in terms of both rank ΔK and slope Δm of the inversion



(a) Unchanged world (b) Inverted world #1 (c) Inverted world #2

Figure 5.26: Comparing appeal and projected final rank. In the unchanged world the results are much more representative of population preferences than in the inverted worlds.

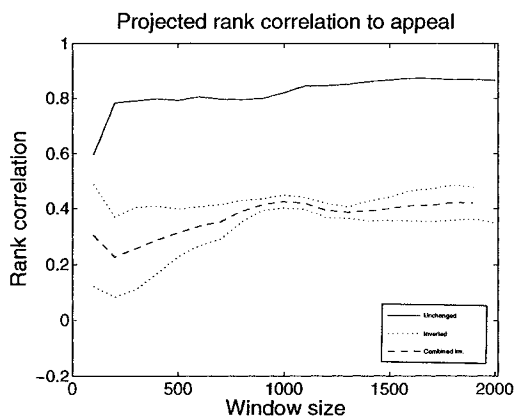
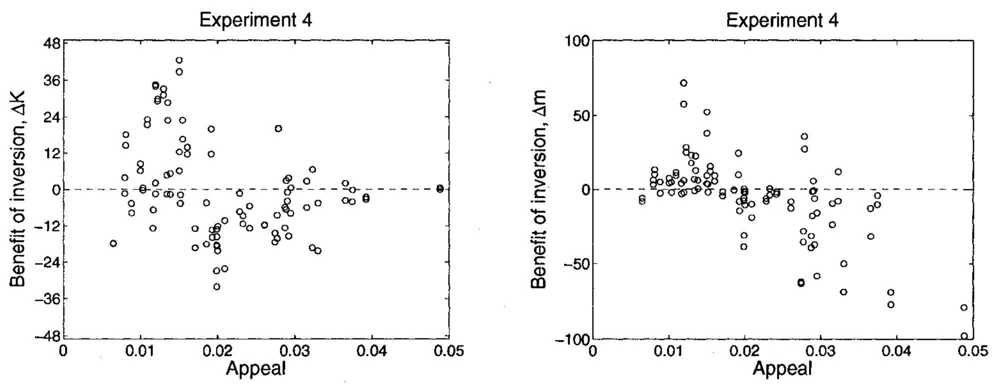


Figure 5.27: Rank correlation between appeal and projected final rank as a function of window size. The results are robust to window size.

as a function of appeal. First, this plot shows that, as suspected, "She Said" was indeed an outlier in terms of appeal. Figure 5.28(a) shows that in terms of projected final ranks, the worst songs benefited, the good songs suffered, and the very best songs were essentially unaffected. In terms of slopes, figure 5.28(b) shows that despite their projected final rank, the best songs were hurt by the inversion.

5.8 Conclusion

Are self-fulfilling prophecies possible in cultural markets? A number of cautions are in order before one should attempt to generalize from these experimental results. Unlike in our experiment, where popularity was manipulated at a single time point in a somewhat extreme manner, manipulation of popularity in the real cultural markets may occur repeatedly, and may also exhibit subtlety and variety. Moreover, whereas our subjects were exposed to only a single source of influence—download counts—information in real cultural markets is far richer than a simple popularity count, including for example reviews (Chevalier and Mayzlin, 2006), and comes from a variety of sources, each of which may have a different impact (Godes et al., 2005). Finally



(a) Benefit of inversion ΔK

(b) Benefit of inversion Δm

Figure 5.28: Benefit of the inversion as a function of appeal. There are two values for each song (one from each inverted world). In general, the best songs were unaffected in terms of rank, but seriously hurt in terms of download rate. Results are jittered to resolve ties.

because our experiment had only 48 songs, the average participant listened to about one-seventh of the music in the market and some listened to almost all he songs. These behaviors are both impossible in real cultural markets where the number of products is overwhelming (Caves, 2000; Vogel, 2004). For this reason we suspect that our finding that the most popular song before the inversion returned to its top spot may not generalize to real cultural markets where if a song is put at the bottom of the popularity ranking (if such a thing exists) it will not generate sufficient listens.

However, we do believe that there are a number of findings from the experiment which may generalize to real cultural markets. On the level of the individual song, it does seem that false information can change the long-run popularity of specific songs and that for some pairs of songs, the inversion in their ordering can become self-sustaining. However, large-scale manipulations are not likely to lock-in because the higher appeal songs partially recover their original positions. The large scale manipulation, however, did decrease the relationship between appeal and success which lead to a decrease in downloads in the inverted worlds. The possibility of self-fulfilling prophecies coupled with the decreased in downloads suggests that the capability of manipulating market information may pose a type of social dilemma (Dawes, 1980)—each band has incentive to manipulate popularity but if too many bands do this it may lead to a reduction in downloads for everyone. One key limitation of this finding is that it only holds if the manipulation of popularity is negatively related to appeal, as was the case in our manipulation.[20]

These findings also suggest a number of additional questions. For example, because we only preformed one type of manipulation on one set of songs, it is not know how our findings would be affected by a more or less severe manipulation or more or less homogeneous songs. Further, it is not known how the results would have been different if the participants had been subjected to stronger or weaker forms of social influence. Ultimately, these questions can only be answered with additional experiments, but here we would like to sketch out some guesses about what these experiments might yield. First, we begin with a simple model which describes how songs gain downloads over time,

$$Pr[d_{jt}] = Pr[l_{jt}] \cdot Pr[d_{jt} \mid l_{jt}] \tag{5.4}$$

where d_{jt} represents a download for song j at time t and l_{jt} is defined equivalently. The first terms it the probability of listen which can be thought of a "social force" and was shown to depend on the popularity rank of the songs (see figure 5.12). The second term can be thought of an intrinsic characteristic of the songs measuring its attractiveness that, after Aizen et al. (2004), we will call "batting average."[21] Thus, the download dynamics of songs can be thought of a combination of the social and intrinsic forces. If we thus consider the relative order of two songs, they can change

[20] A mathematical model of information manipulation which reaches similar conclusions is presented in Dellarocas (2006).

[21] For the bulk of the anylysis in the dissertation, we chose our measure of appeal, market share in the independent world, over the batting average because, by implicitly including a measure of the attractiveness of the songs names, it made for a more fair comparison between the social influence and independent conditions.

order if the extra listens that the leading song gets are offset by the high batting average of the trailing song. Thus, we suspect that if either the social influence were stronger or the songs were more homogeneous in terms of appeal we would see more self-fulfilling prophecies. Also, we suspect that if there were more songs we would also see more self fulfilling prophecies because no matter how attractive a song ($Pr[d_j \mid l_j]$) it can't gain downloads if it doesn't get listens. Ultimately, these intuitions should be formalized using a mathematical model, probably based on equation 5.4.

Beyond the study of cultural markets, we hope the experimental design used here (figure 5.3) will be helpful for others interested in the study path-dependent systems (Arthur, 1994; Liebowitz and Margolis, 1995; Mahoney, 2000; Pierson, 2000; Page, 2006). These systems, in which a single event can permanently alter system trajectory, are thought to be common in social science, but they are difficult to study using observational data. Our experimental design offers two advantages. First, the "multiple worlds" feature allows counterfactual outcomes to be directly compared. Second, both the individual-level origins of path-dependence (in this case social influence) and the collective-level consequences (popularity) can be measure simultaneously and their interaction can be observed over time (Schelling, 1978; Coleman, 1990; Hedström, 2005).

The four experiments just described provide strong support for the proposed model. In the next chapter we turn our attention to other questions that can be addresses with these data.

Chapter 6

Other issues

The previous chapters have answered questions most directly related to the puzzle that motivated this dissertation. In this chapter we use the data that was collected to also address other related questions.

6.1 Was everything determined by the first few subjects?

One of the main conclusions of the lock-in, path dependence, and information cascades literature is that the first few participants have a disproportionate effect on system dynamics (Arthur et al., 1987; Arthur, 1989; Banerjee, 1992; Bikhchandani et al., 1992; Liebowitz and Margolis, 1995; Anderson and Holt, 1997; Bikhchandani et al., 1998; Mahoney, 2000; Page, 2006). We can see if this was the case in our data.[1]

If the early participants had a disproportionate impact on the dynamics then we would expect that the final state of the system will be highly correlated with the very early states of the system. The correlation between the initial state and final state is 0 (by definition) and the correlation between the final state and the final

[1] In some sense this question is intimately linked with the issues explored in experiment 4.

state is 1 (again, by definition). However, the time-series of the correlation between the current state and the final state—the "correlation profile"—gives us important information about lock-in.

To be clear, let $\vec{d_t}$ be a vector of download counts at time t and let $\vec{R_t}$ be the ranking of the songs also at time t. It is the case that,

$$r_0 = corr(\vec{d_0}, \vec{d_T}) = 0 \tag{6.1}$$

$$\rho_0 = corr(\vec{R_0}, \vec{R_T}) = 0 \tag{6.2}$$

where r is the correlation, ρ is the rank correlation, and T is the final time. It is also the case that

$$r_T = corr(\vec{d_T}, \vec{d_T}) = 1 \tag{6.3}$$

$$\rho_T = corr(\vec{R_T}, \vec{R_T}) = 1 \tag{6.4}$$

What we are interested in is how r_t and ρ_t change as the worlds progress. If the first subject completely determined the outcome, the ultimate lock-in, then $\rho_1 = 1, \rho_2 = 1, \ldots \rho_T = 1$. As we will see, however, this was not the case.

6.1.1 Rank correlation profile, ρ_t

The rank correlation profiles from experiments 1, 2 and 3 are plotted in figures 6.1, 6.2, and 6.3.[2] In all worlds, the results are far short of complete lock-in, that is, ρ_t does not approach 1 very quickly, but the rank correlation profiles are generally increasing over time.

One way to get a sense of the speed to the final state is to compare to the rank correlation profile in the independent world. However, this introduces some complications as we will now describe. The independent world had about twice as

[2]The rank correlation profile and correlation profile from experiment 4 are not of interest here because of the inversion.

many participants as the social influence worlds so in figures 6.1, 6.2, and 6.3 we have set T for the independent world to be the number of subjects that were in the social influence world to which the comparison is made. For example, for experiment 1, the independent condition rank correlation profile for world 1 describes the approach to the final state of the independent world after 702 subjects; for world 2, it describes the approach to the state of the independent world after 699 subjects. When comparing the rank correlation profiles for the independent and social influence worlds, lock-in would be indicated if the social influence world moved faster to its final state than the independent world (i.e., if the solid line was above the dashed line in the figures). However, inspection of the figures reveals that this does not seem to be strongly the case—sometimes the social influence world approaches faster, but sometimes the independent world does.

One possible reason that these figures fail to show strong lock-in is that the comparison to the independent world may not be a fair one because the composition of the groups are different; the social influence worlds have participants recruited during the entire experiment, but the independent world here only has participants from the first half. To account for this problem we also created figures 6.4, 6.5, and 6.6 which plot the entire independent world against a social influence world with the common x-scale being proportion of subjects. This approach ensures that the composition of the groups being compared are the same, but it is not without problem either because it involves comparing outcomes based on different numbers of subjects. For example, when the system is at 50% of subjects, in the influence world this represents about 350 people but in the independent world it represents about 700 people. These plots which have the proportion of subjects on the x-axis (6.4, 6.5, and 6.6) are not qualitatively different from those which have the absolute number of subjects on the x-axis (6.1, 6.2, and 6.3), so the choice of x-axis does not explain the lack of lock-in.

To summarize, what all of these figures show is that the approach to the final

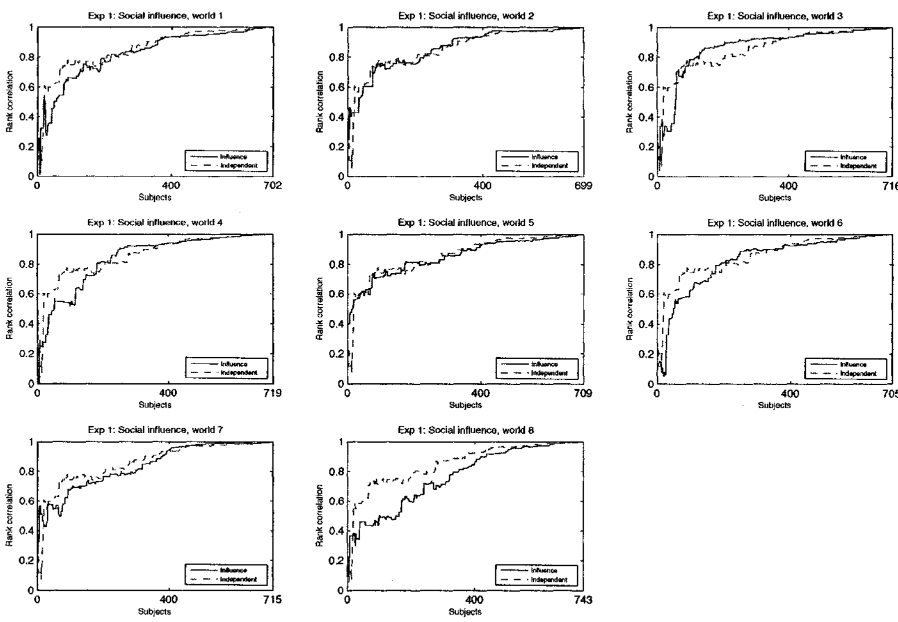


Figure 6.1: Rank correlation profile in all social influence worlds in experiment 1 (x-axis = number of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

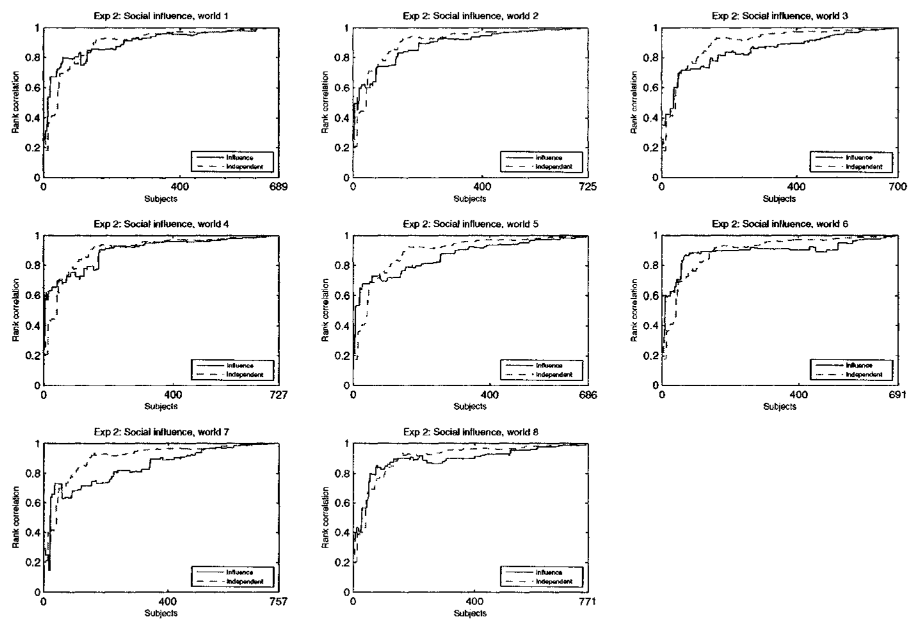


Figure 6.2: Rank correlation profile in all social influence worlds in experiment 2 (x-axis = number of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

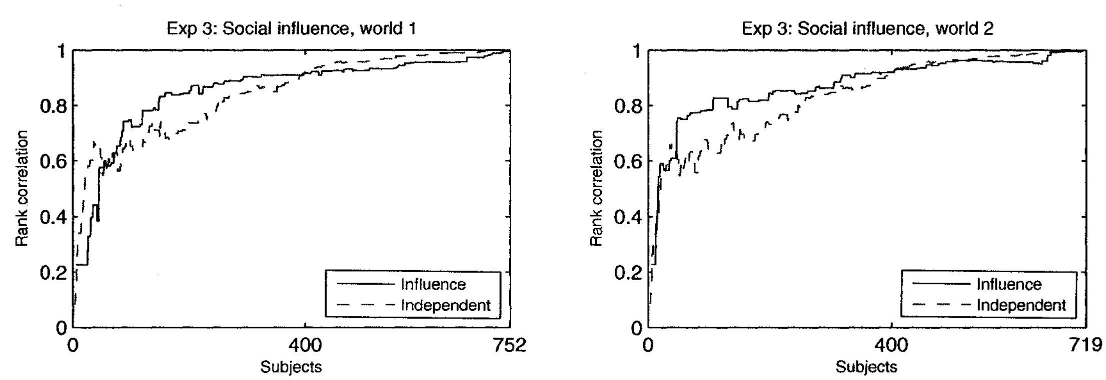


Figure 6.3: Rank correlation profile in all social influence worlds in experiment 3 (x-axis = number of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

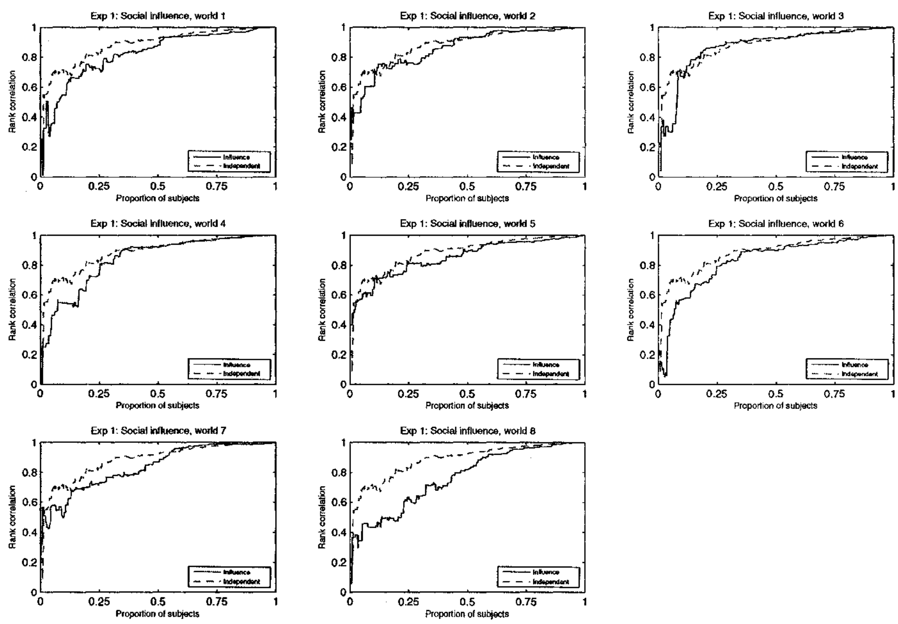


Figure 6.4: Rank correlation profile in all social influence worlds in experiment 1 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

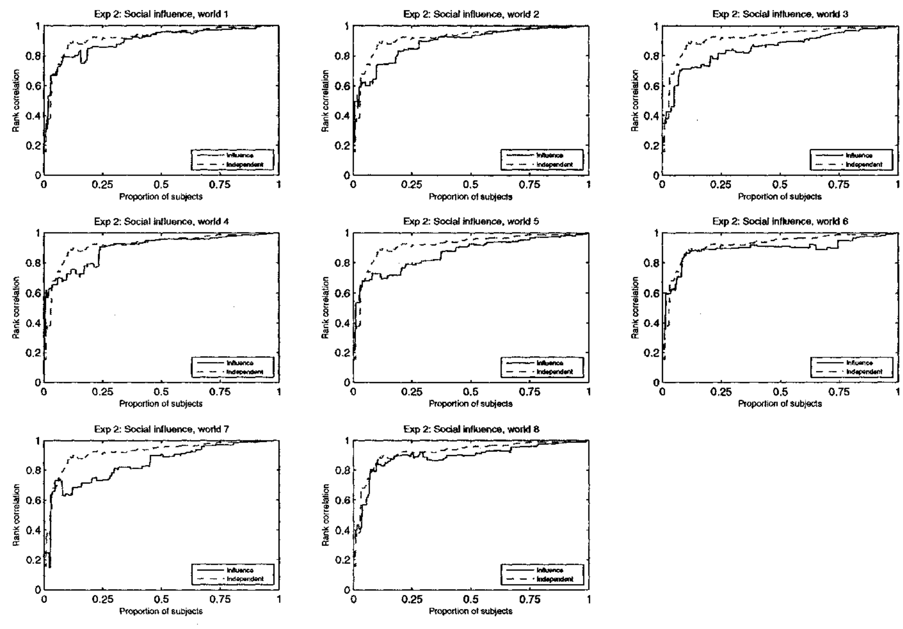


Figure 6.5: Rank correlation profile in all social influence worlds in experiment 2 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

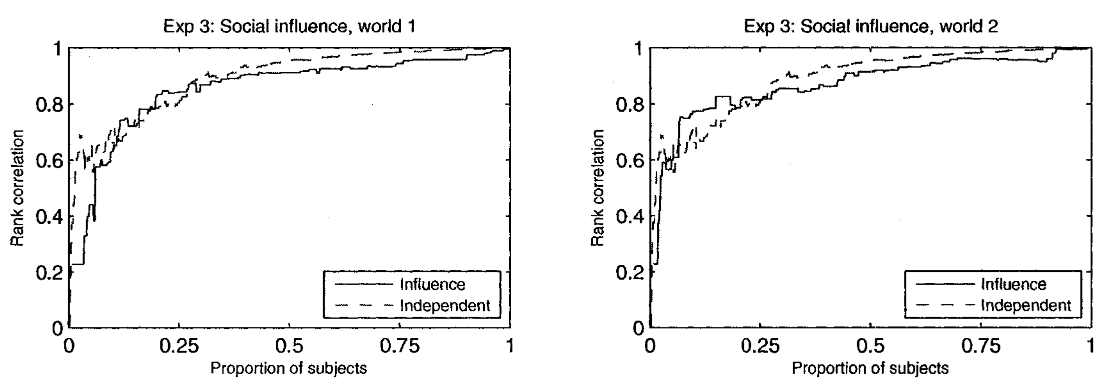


Figure 6.6: Rank correlation profile in all social influence worlds in experiment 3 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

state happened at a similar rate in the independent and social influence worlds. This seems to argue against a strong form of lock-in or any kind of sensitive dependence on initial conditions. However, this conclusion, which is based on the rank correlation profile, seems to be slightly contradicted by analysis of the correlation profile as we will see in the next section.

6.1.2 Correlation profile, r_t

In addition to the rank correlation profile, ρ_t, we can also calculate the correlation profile, r_t. These figures (6.7, 6.8, and 6.9) seem to indicate stronger lock-in. The reason for the difference between the correlation profile and rank correlation profile is currently unknown, but one possible explanation is that rank correlation is particularly susceptible to noise at the bottom of the popularity distribution where a few downloads can greatly change the rank of a song.

Also, as with the rank correlation profile, we plotted the results with the x-axis as the proportion of subjects (figures 6.10, 6.11, and 6.12). This makes the lock-in look slightly weaker, especially for experiment 2 (compare figure 6.8 and figure 6.11), but we don't yet know why this is the case.

6.1.3 Correlation profile for the top 10 songs

The previous sections have found relatively weak evidence of lock-in. One possible reason is that there is lots of noise at the bottom of the popularity distribution which is masking lock-in for the most popular songs. Therefore, in figures 6.13, 6.14, and 6.15 we track the correlation profile, r_t, of the songs which ended up in the top 10 in each world.[3] To ensure a fair comparison, this correlation profile is then

[3] To be clear, these 10 songs could be different in the different social influence worlds. To see which songs are in the top 10 in each world, see appendix 3 (for experiments 1 and 2) and appendix 4 (for experiment 3).

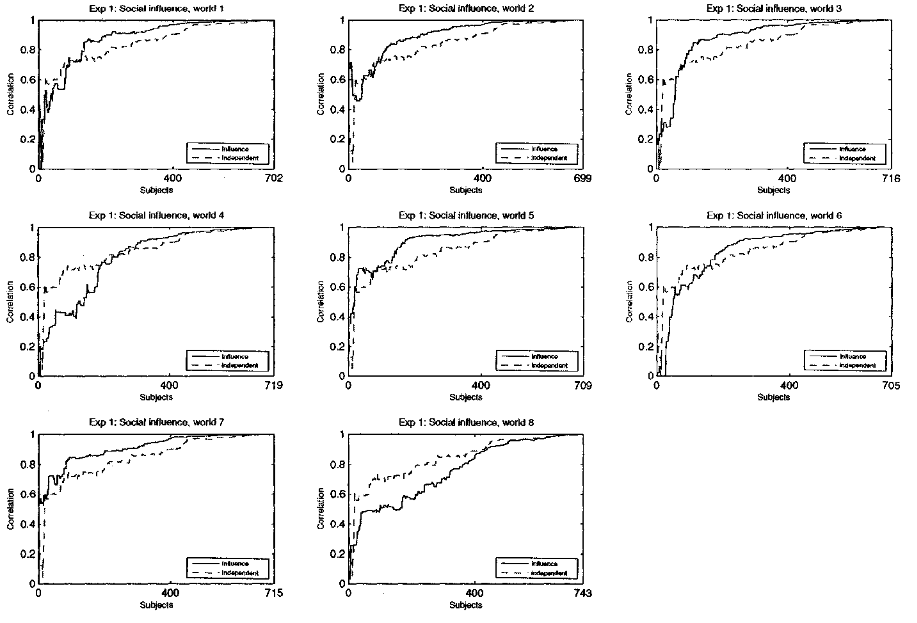


Figure 6.7: Correlation profile in all social influence worlds in experiment 1 (x-axis = number of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

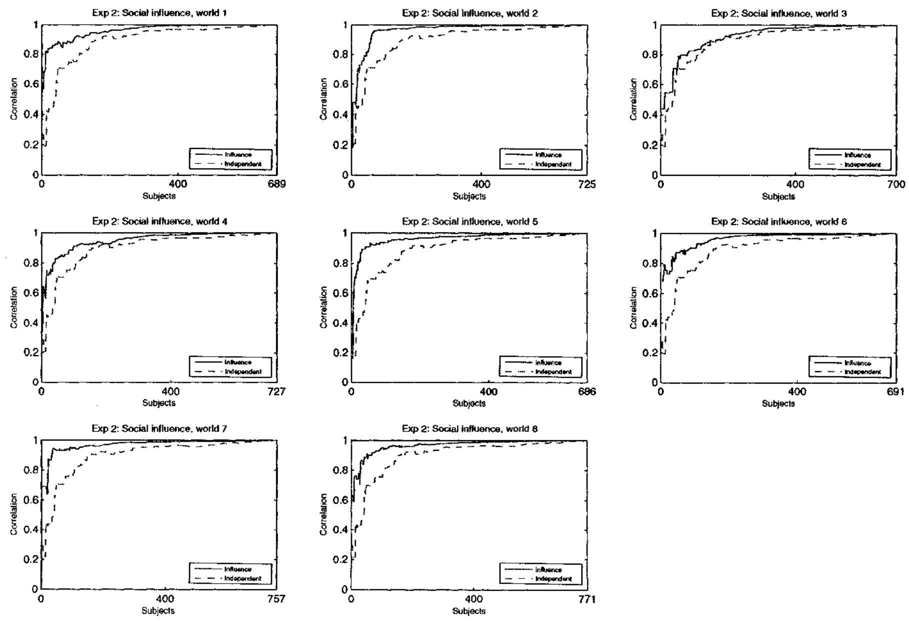


Figure 6.8: Correlation profile in all social influence worlds in experiment 2 (x-axis = number of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

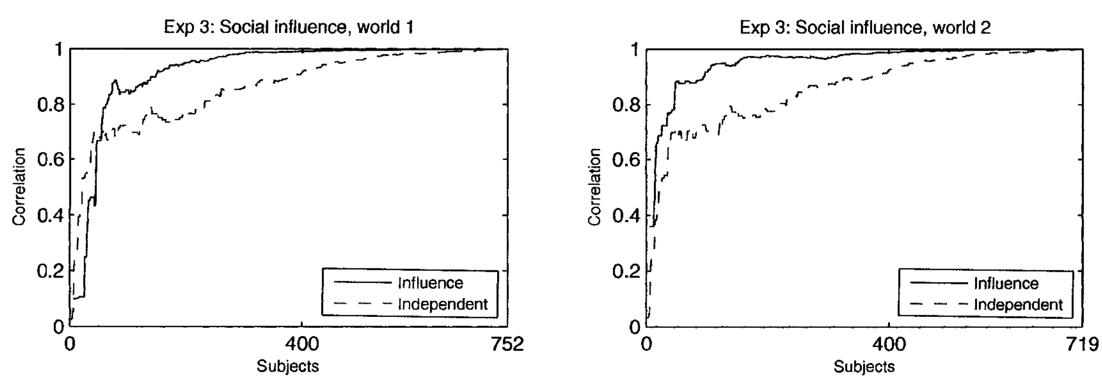


Figure 6.9: Correlation profile in all social influence worlds in experiment 3 (x-axis = number of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

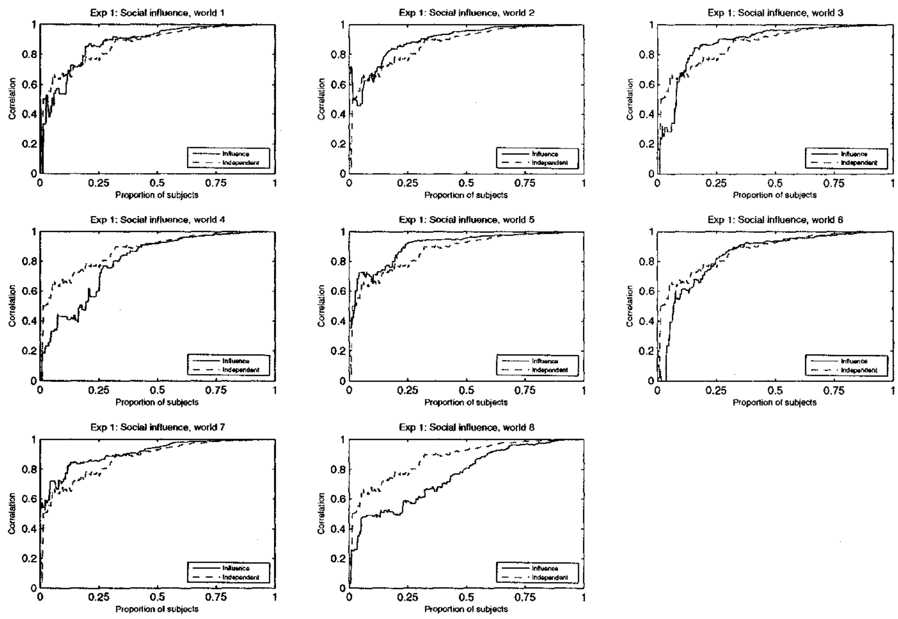


Figure 6.10: Correlation profile in all social influence worlds in experiment 1 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

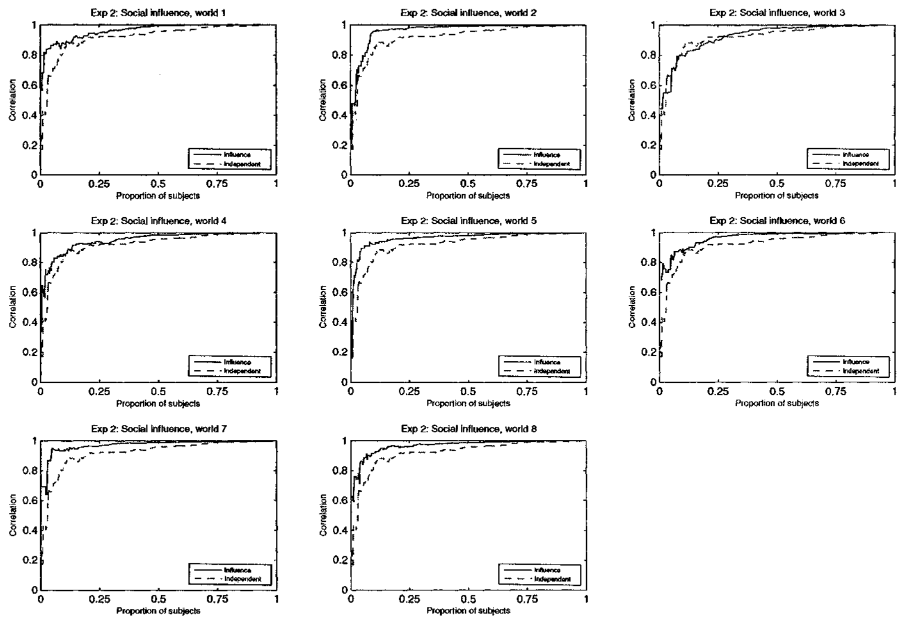


Figure 6.11: Correlation profile in all social influence worlds in experiment 2 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

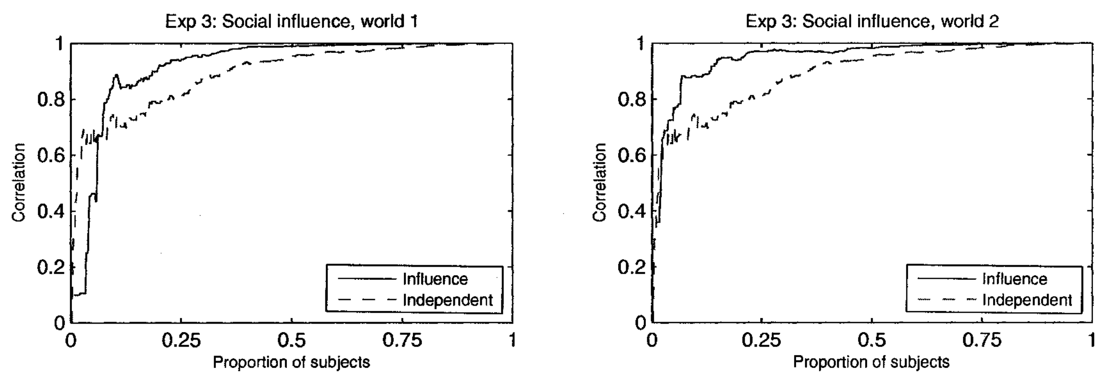


Figure 6.12: Correlation profile in all social influence worlds in experiment 3 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

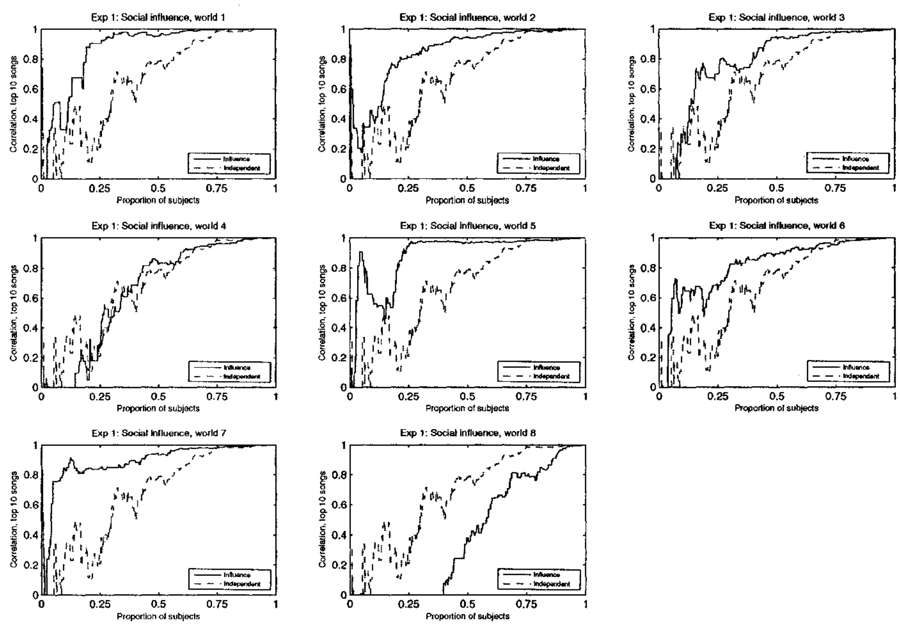


Figure 6.13: Correlation profile for the top 10 songs in each world in experiment 1 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

compared against that of the top 10 songs independent condition.

For ease of presentation we will only present plots for the top 10 songs where the x-axis the proportion of subjects. Plots for the correlation profiles of the top 5 songs, rank correlation profiles, and plots with different x-axis values are presented in appendix E. The general conclusion that can be drawn from figures 6.13, 6.14, and 6.15 is that there is evidence of lock-in for the top 10 songs.

6.1.4 Market share dynamics

Yet another approach to the question of lock-in simply involves plotting the market shares for the songs over time. Whenever two songs change rank (i.e. their market shares trajectories cross) that indicates lack of lock-in. The market share

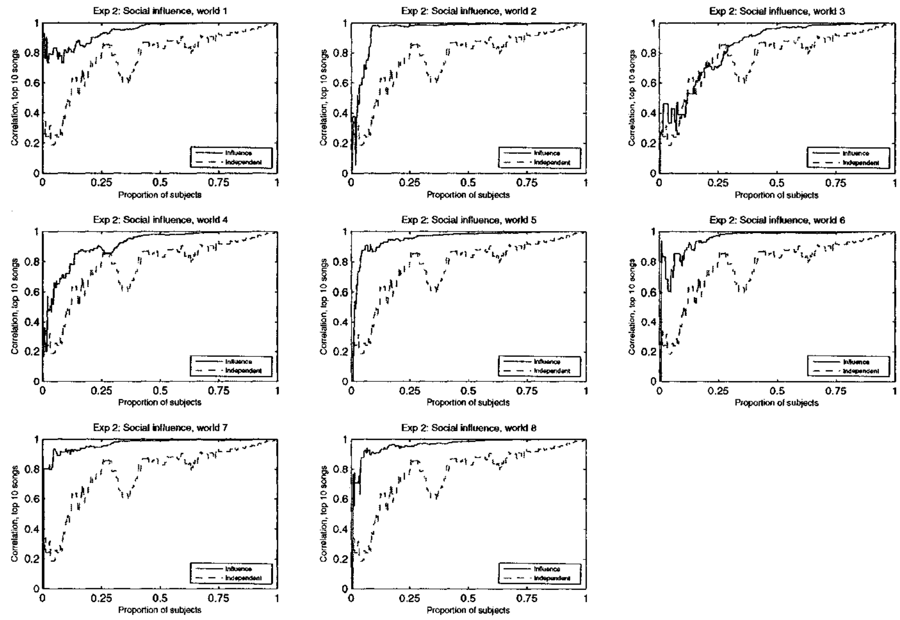


Figure 6.14: Correlation profile for the top 10 songs in each world in experiment 2 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

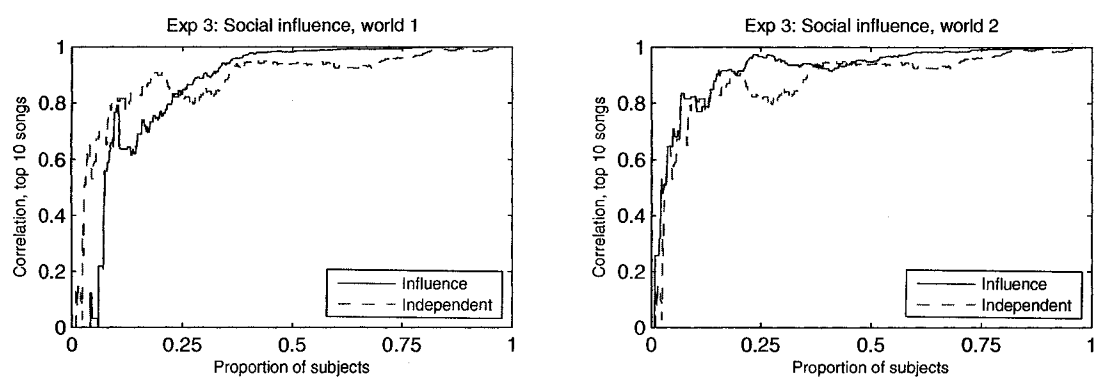


Figure 6.15: Correlation profile for the top 10 songs in each world in experiment 3 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

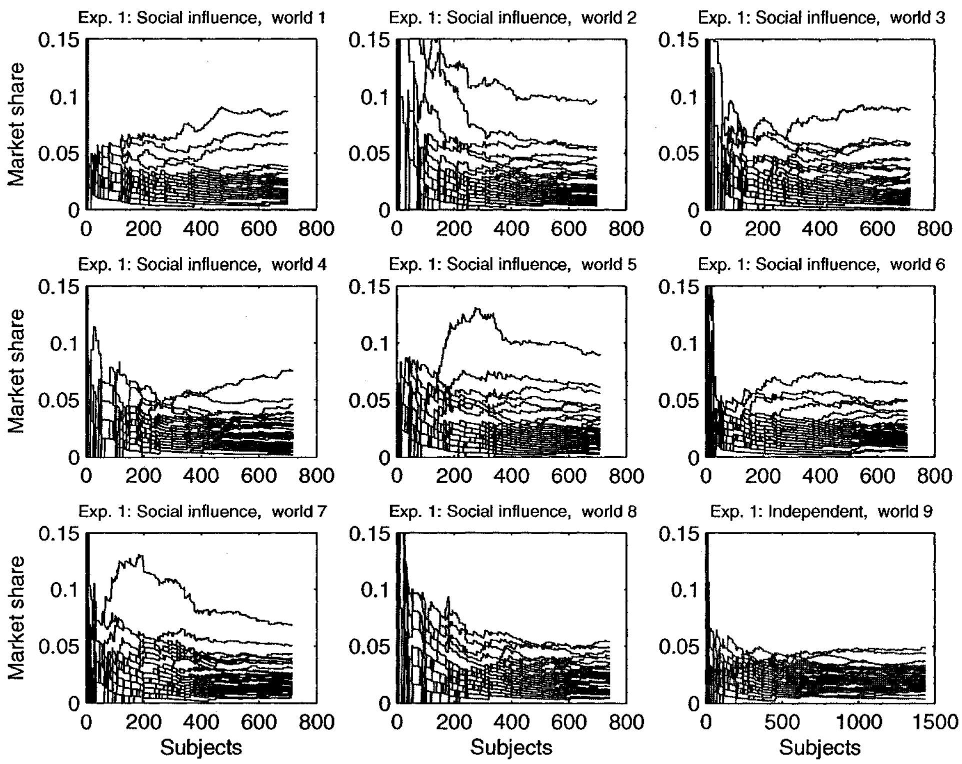


Figure 6.16: Dynamics of the market share of all songs in all worlds in experiment 1.

trajectories are presented in figures 6.16, 6.17, and 6.18. As was indicted by our correlation profiles, it seems that there are some order changes, which argues against a strong form of lock-in. However, it is hard to know how to summarize these figures numerically.

6.1.5 Next steps and conclusions

In conclusion, it seems that there was some evidence of lock-in for the top 10 songs, but the evidence of lock-in for the entire system is weak. There are a number of

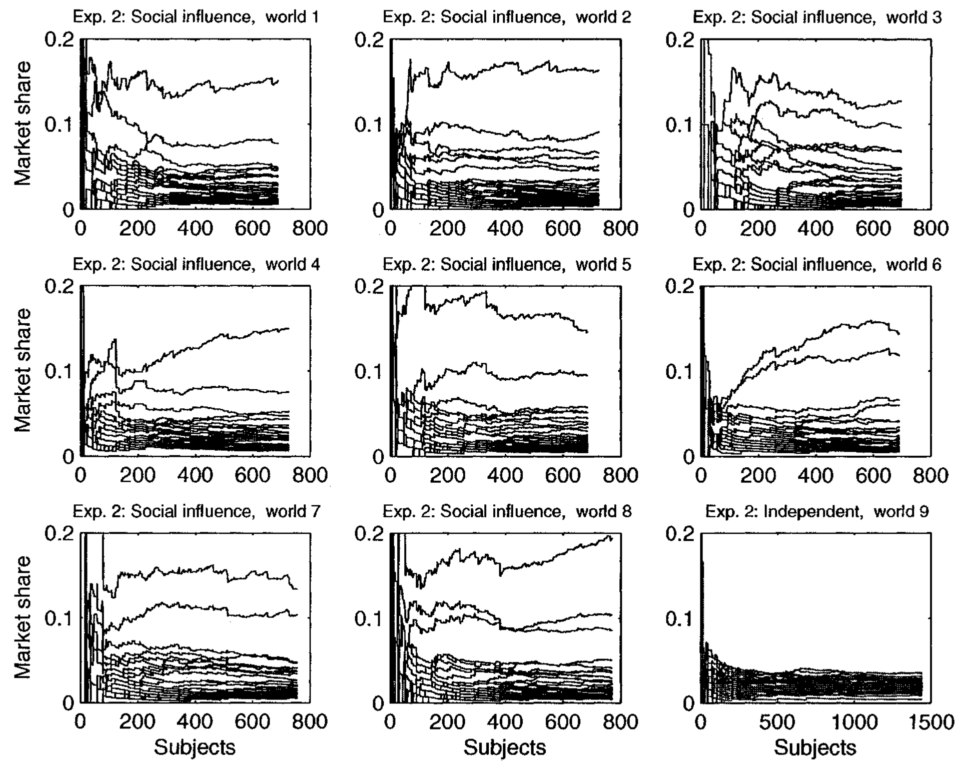


Figure 6.17: Dynamics of the market share of all songs in all worlds in experiment 2.

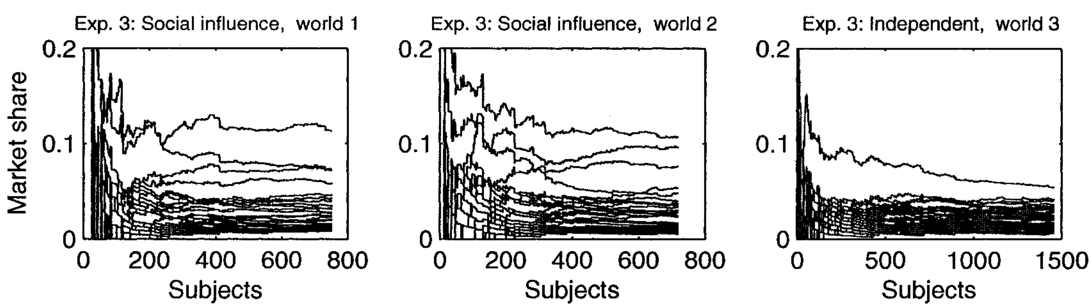


Figure 6.18: Dynamics of the market share of all songs all worlds in experiment 3.

possible next steps in this area including exploring alternative definitions of lock-in.[4] Additionally, we have not made use of the multiple-world nature of our design; that is, we have not compared how the worlds diverge and perhaps this should be considered as a test for lock-in. An additional analysis could involve comparing rates of lock-in across experiments. For example, we might suspect that because the social influence was stronger in experiment 2, the correlation profiles from experiment 2 would look different from experiment 1. An additional important question is whether we can use some of these ideas to test for lock-in in observation data.

[4]The conception of lock-in used here is a system-level one. It is also the case that some scholars consider lock-in at the individual-level. For example, Page (2006) writes, "lock-in means that one choice or action becomes better than any other one because a sufficient number of people have already made that choice." We do not feel that this individual-level approach is the most appropriate for the questions here.

6.2 Were some band names and song names better than others?

The listen count in the independent world serves as a measure of the attractiveness of the band name and song name. Tables 6.1, 6.2, 6.3, and 6.4 report these listen counts. In experiments 1 and 2, the most listened to song had about twice as many listens as the least listen to song. In experiments 3 and 4, this difference was smaller.

6.2.1 Were the songs with more attractive names downloaded at a higher rate?

A related question is whether the attractiveness of the song/band name might have affected subjects' download decision. The idea is that since people are first exposed to the song/band name, this might predispose them to download the song. To check this we calculated the correlation between the listens in the independent condition (a measure of song/band name attractiveness) and batting average in the independent condition. These correlations were low, but positive, in experiments 1 and 2 ($r = 0.19$ and $r = 0.12$) and basically 0 in experiments 3 and 4 ($r = -0.02$ and $r = -0.03$). We would not interpret this is a strong evidence that the song/band name influenced subjects download decisions. Results are presented graphically in figure 6.19.

Band name	Song name	Listens
stunt monkey	inside out	177
the fastlane	til death do us part (i dont)	168
ember sky	this upcoming winter	164
miss october	pink agression	164
by november	if i could take you	157
parker theory	she said	141
benefit of a doubt	run away	138
the broken promise	the end in friend	137
deep enough to die	for the sky	135
silent film	all i have to say	135
up falls down	a brighter burning star	134
hartsfield	enough is enough	124
beerbong	father to son	124
moral hazard	waste of my life	122
stranger	one drop	122
hall of fame	best mistakes	120
dante	lifes mystery	118
star climber	tell me	117
silverfox	gnaw	113
52metro	lockdown	113
unknown citizens	falling over	111
not for scholars	as seasons change	111
far from known	route 9	107
the thrift syndicate	2003 a tragedy	105
a blinding silence	miseries and miracles	104
summerswasted	a plan behind destruction	103
simply waiting	went with the count	102
nooner at nine	walk away	102
fading through	wish me luck	101
sum rana	the bolshevik boogie	101
post break tragedy	florence	100
hydraulic sandwich	separation anxiety	100
salute the dawn	i am error	99
undo	while the world passes	98
go mordecai	it does what its told	96
this new dawn	the belief above the answer	96
evan gold	robert downey jr	96
selsius	stars of the city	94
shipwreck union	out of the woods	94
art of kanly	seductive intro, melodic breakdown	90
the calefaction	trapped in an orange peel	90
forthfading	fear	90
sibrian	eye patch	90
up for nothing	in sight of	82
secretary	keep your eyes on the ballistics	82
ryan essmaker	detour (be still)	81
drawn in the sky	tap the ride	74
cape renewal	baseball warlock v1	72

Table 6.1: The number of listens for each song in the independent world of experiment 1.

Band name	Song name	Listens
the broken promise	the end in friend	164
by november	if i could take you	162
stunt monkey	inside out	159
the fastlane	til death do us part (i dont)	149
hall of fame	best mistakes	148
miss october	pink agression	146
parker theory	she said	143
ember sky	this upcoming winter	138
up falls down	a brighter burning star	137
post break tragedy	florence	134
dante	lifes mystery	133
silent film	all i have to say	131
deep enough to die	for the sky	129
go mordecai	it does what its told	126
nooner at nine	walk away	126
undo	while the world passes	120
moral hazard	waste of my life	120
unknown citizens	falling over	120
hartsfield	enough is enough	118
beerbong	father to son	117
benefit of a doubt	run away	116
a blinding silence	miseries and miracles	116
art of kanly	seductive intro, melodic breakdown	114
stranger	one drop	114
star climber	tell me	113
the thrift syndicate	2003 a tragedy	112
52metro	lockdown	111
not for scholars	as seasons change	109
forthfading	fear	109
summerswasted	a plan behind destruction	108
this new dawn	the belief above the answer	107
up for nothing	in sight of	106
hydraulic sandwich	separation anxiety	106
fading through	wish me luck	105
the calefaction	trapped in an orange peel	105
simply waiting	went with the count	105
selsius	stars of the city	104
silverfox	gnaw	102
drawn in the sky	tap the ride	102
ryan essmaker	detour (be still)	100
far from known	route 9	99
salute the dawn	i am error	98
evan gold	robert downey jr	98
shipwreck union	out of the woods	98
sum rana	the bolshevik boogie	96
sibrian	eye patch	95
secretary	keep your eyes on the ballistics	89
cape renewal	baseball warlock v1	86

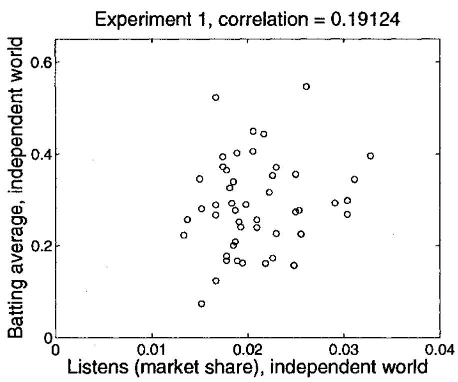
Table 6.2: The number of listens for each song in the independent world in experiment 2.

Band name	Song name	Listens
go mordecai	it does what its told	343
ember sky	this upcoming winter	337
miss october	pink agression	321
post break tragedy	florence	291
sum rana	the bolshevik boogie	291
dante	lifes mystery	285
evan gold	robert downey jr	277
hall of fame	best mistakes	276
stunt monkey	inside out	276
parker theory	she said	275
deep enough to die	for the sky	274
silent film	all i have to say	268
by november	if i could take you	261
art of kanly	seductive intro, melodic breakdown	257
beerbong	father to son	253
undo	while the world passes	252
silverfox	gnaw	251
up for nothing	in sight of	248
the calefaction	trapped in an orange peel	248
hydraulic sandwich	separation anxiety	243
the fastlane	til death do us part (i dont)	242
benefit of a doubt	run away	242
the thrift syndicate	2003 a tragedy	240
52metro	lockdown	239
stranger	one drop	239
cape renewal	baseball warlock v1	238
star climber	tell me	236
the broken promise	the end in friend	236
up falls down	a brighter burning star	234
nooner at nine	walk away	234
shipwreck union	out of the woods	234
ryan essmaker	detour (be still)	232
not for scholars	as seasons change	231
secretary	keep your eyes on the ballistics	231
this new dawn	the belief above the answer	230
unknown citizens	falling over	227
selsius	stars of the city	225
hartsfield	enough is enough	223
moral hazard	waste of my life	222
far from known	route 9	222
salute the dawn	i am error	221
fading through	wish me luck	213
summerswasted	a plan behind destruction	212
sibrian	eye patch	211
a blinding silence	miseries and miracles	208
drawn in the sky	tap the ride	206
forthfading	fear	199
simply waiting	went with the count	190

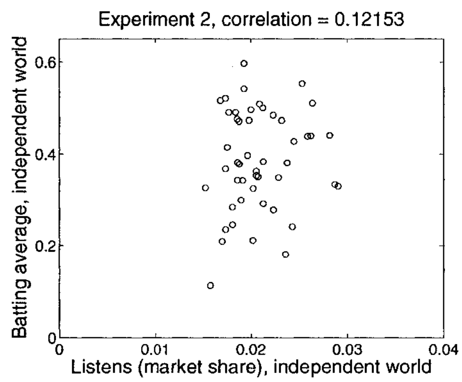
Table 6.3: The number of listens for each song in the independent world in experiment 3.

Band name	Song name	Listens
ember sky	this upcoming winter	847
go mordecai	it does what its told	771
miss october	pink agression	768
deep enough to die	for the sky	741
parker theory	she said	739
sum rana	the bolshevik boogie	732
stunt monkey	inside out	716
silent film	all i have to say	710
evan gold	robert downey jr	704
post break tragedy	florence	695
by november	if i could take you	690
dante	lifes mystery	690
hall of fame	best mistakes	689
the fastlane	til death do us part (i dont)	676
art of kanly	seductive intro, melodic breakdown	673
this new dawn	the belief above the answer	654
silverfox	gnaw	651
the broken promise	the end in friend	644
up falls down	a brighter burning star	632
beerbong	father to son	632
undo	while the world passes	629
the thrift syndicate	2003 a tragedy	609
the calefaction	trapped in an orange peel	608
up for nothing	in sight of	602
a blinding silence	miseries and miracles	601
fading through	wish me luck	597
star climber	tell me	596
not for scholars	as seasons change	596
52metro	lockdown	593
sibrian	eye patch	593
shipwreck union	out of the woods	590
nooner at nine	walk away	588
hartsfield	enough is enough	587
benefit of a doubt	run away	583
hydraulic sandwich	separation anxiety	580
selsius	stars of the city	577
secretary	keep your eyes on the ballistics	568
unknown citizens	falling over	565
cape renewal	baseball warlock v1	564
far from known	route 9	564
ryan essmaker	detour (be still)	564
moral hazard	waste of my life	560
stranger	one drop	555
salute the dawn	i am error	546
forthfading	fear	521
summerswasted	a plan behind destruction	520
drawn in the sky	tap the ride	517
simply waiting	went with the count	515

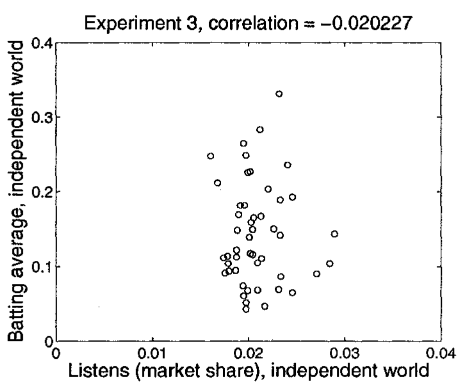
Table 6.4: The number of listens for each song in the independent world in experiment 4.



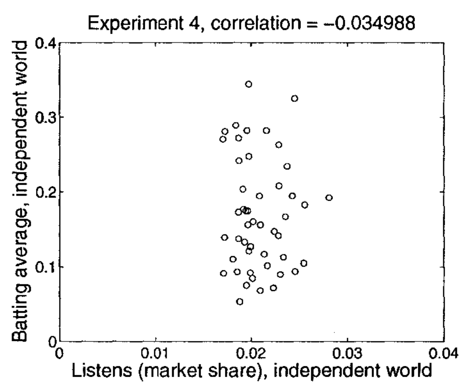
(a) Experiment 1



(b) Experiment 2



(c) Experiment 3



(d) Experiment 4

Figure 6.19: Comparing the attractiveness of the song/band name and the batting average. In experiments 1 and 2 there was a weak relationship, but in experiments 3 and 4 there was no relationship. From this data we would not conclude that there is strong evidence that the attractiveness of the song/band name influenced subjects' download decisions.

6.3 How did screen location affect listen decisions?

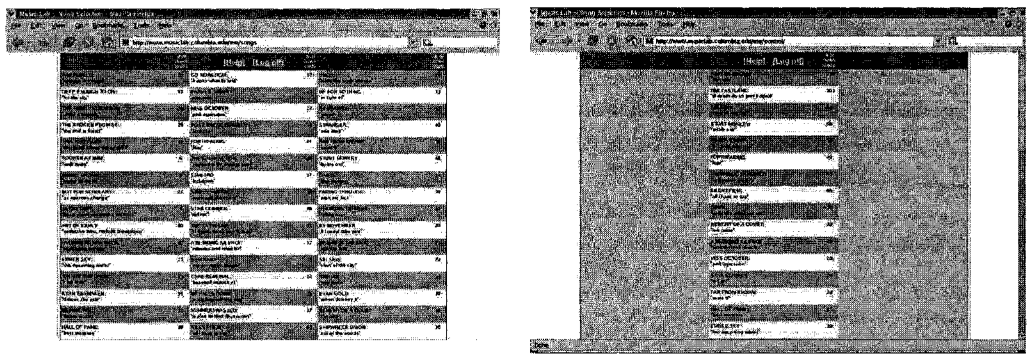
Within the web-design community there is a belief that the location of content on a webpage affects the allocation of user attention. For example, content "above the fold," that is, visible without scrolling, is thought to receive more attention than content below the fold. In our experiments we can explore the relationship between screen location and listens. First, we will try to understand how screen location affects listen decisions in the absence of any information about the popularity of the songs. Then we will examine if popularity affected participants' listen decisions, above and beyond its effect on screen location (recall that in experiments 2, 3, and 4 popular songs were given better screen locations).

6.3.1 Independent world

In experiment 1 the songs were presented to the participants in a 16×3 grid (figure 6.20(a)) and in experiments 2, 3, and 4 they were presented in a 1-column list (figure 6.20(b)). By focusing on the independent worlds we can isolate a pure screen location effect because in these worlds the songs were randomly assigned to screen locations for each participant.

Figure 6.21 plots the number of listens for each screen location in the independent world in experiment 1. By far the most listened to location was the upper-left-hand corner. Further, many subjects appeared to have read down the first column rather than across the top row. In general, for a given row, more listens went to column 1, then column 2, and finally column 3. One exception to this pattern is the last row in which column 3 had the most listens probably because row 16, column 3 is the lower right hand corner, a location that is highlighted by the grid layout.

Similarly, figure 6.22 plots the number of listens for each screen location in the independent world in experiments 2, 3, and 4. Generally, locations higher on the



(a) Experiment 1 (b) Experiment 2, 3, and 4

Figure 6.20: Screenshots of the song menus. In experiment 1 the songs were presented to the participants in a 16×3 grid and in experiments 2, 3, and 4 they were presented in a 1-column list.

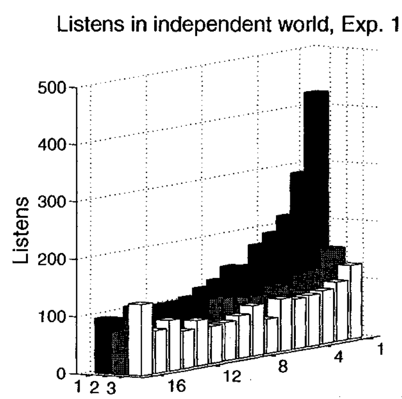
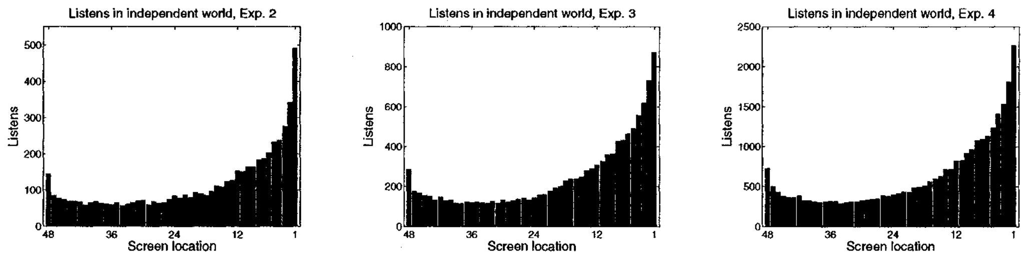


Figure 6.21: Listens by screen location in the independent world in experiment 1. In general, the higher and more to the left the location, the more listens.



(a) Experiment 2 (b) Experiment 3 (c) Experiment 4

Figure 6.22: Listens by screen location in the independent world in experiments 2, 3, and 4. In all three experiments, locations higher on the list earned more listens with the exception of a slight increase for songs at the very bottom of the list.

list earned more listens with the exception of a slight increase for songs at the very bottom of the list. Further, these results are similar across experiments, even thought the participants in experiment 2 were quite different from those in experiment 3 and 4.

In summary, subjects in the independent worlds were clearly influenced by screen location in their listening decisions. This finding then raises the question of how much of the observed effect of popularity on listening decision was actually mediated by the screen location. For this we require comparing listing patterns in the independent and social influence world.

6.3.2 Social influence world

In the previous chapters of this dissertation we have shown that subjects in the social influence worlds were more likely to listen to songs that were more popular. However, the findings in the previous section suggest that this pattern may be at least partially determined by screen location, rather than popularity.[5] This naturally raises the question, did the download counts alter subjects' listening choices above

[5]To be clear, for the substantive questions of interest we don't think it matters exactly why participants in our experiment choose to listen to the more popular songs.

and beyond their effect screen location? In other words, was the most popular song an influence world (where the download counts were visible) more likely to be listened to than whatever song was randomly placed at the top of the screen in the independent world (where the download counts are not visible)?

To address this question, for each experiment, we calculated the proportion of listens that went to each screen location in each world. More specifically, let $l_{k,w}$ be number of listens to the song in screen location k in world w. Then, we calculated,

$$\lambda_{k,w} = \frac{l_{k,w}}{\sum_{k=1}^{k=48} l_{k,w}} \tag{6.5}$$

Next, we calculated

$$\Delta\lambda_{k,w} = \lambda_{k,w} - \lambda_{k,ind} \tag{6.6}$$

which measure the difference in proportion of listens for a given location in a social influence and independent world. A positive value of $\Delta\lambda_{k,w}$ means that the location was more listened to in the social influence world (relative to other locations in the social influence world) than in the independent world (relative to other locations in the independent world). Note that by definition it is the case that for a given world, $\sum_k \Delta\lambda_{k,w} = 1$.

To build intuition, figure 6.23 plots $\lambda_{k,w}$ for the four worlds in experiment 4.[6] Next figure 6.24 plots $\Delta\lambda_{k,w}$ for the three social influence worlds. These plots show that in the social influence worlds the top and bottom locations get a higher proportion of listens than they do in the independent world.[7] In other words, it seems that the presence of the download counts triggers a "ranking heuristic" which causes

[6]One thing that does not appear to affect these results is the distinction between songs that are above and below the fold. As a point of reference, on my computer, 15 songs were visible without scrolling and there is no sharp line at screen location 15. Of course, screen have different resolutions so the dividing line for above and below the fold will vary by computer.

[7]One unusual pattern in figure 6.24(a) is the non-monotonic behavior around screen location 3 to 6. However, this pattern is also visible in the raw data from the unchanged world in figure 6.23. We do not know this origin of this anomaly, but we do not think it is cause for concern.

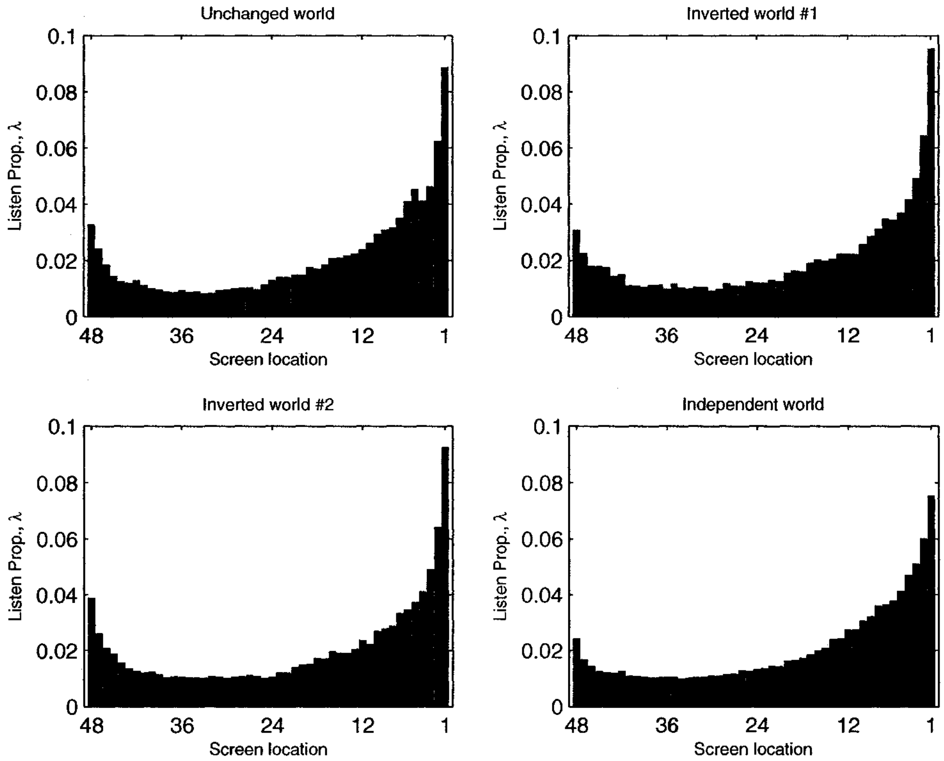
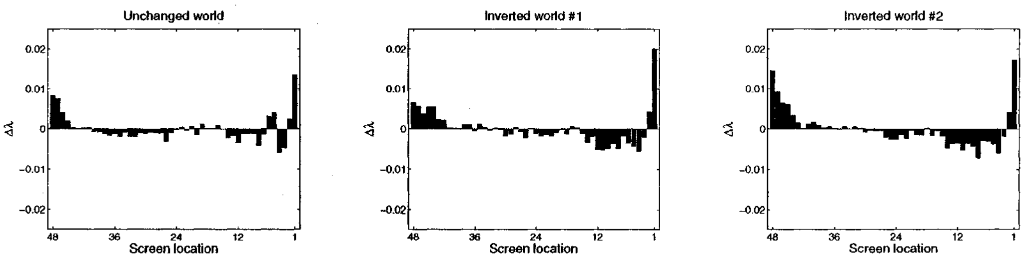


Figure 6.23: Proportion of listens by screen location in all four worlds from experiment 4.

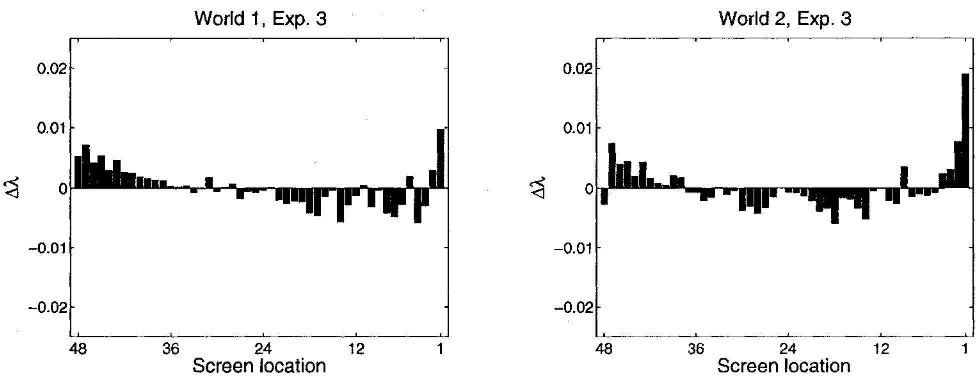
some participants to prefer the top or bottom location (relative to other locations). Similar results are observed for experiment 3 (figure 6.25). Results from experiment 2 are also similar except that in this experiment subjects did not seem to favor the least popular songs (figures 6.26). We don't know the reason for this difference.

In summary, screen location seems to have played a strong role in determining which songs participants choose to listen to in the independent condition. However, providing subjects with information on the popularity of the songs had an effect above and beyond screen location.



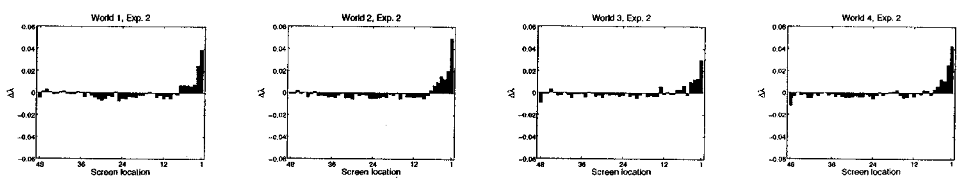
(a) Unchanged world (b) Inverted world #1 (c) Inverted world #2

Figure 6.24: $\Delta\lambda_{k,w}$ for the three social influence worlds in experiment 4. In all three social influence worlds participants were relatively more likely to listen to songs at the top and bottom screen locations compared to participants in the independent world.

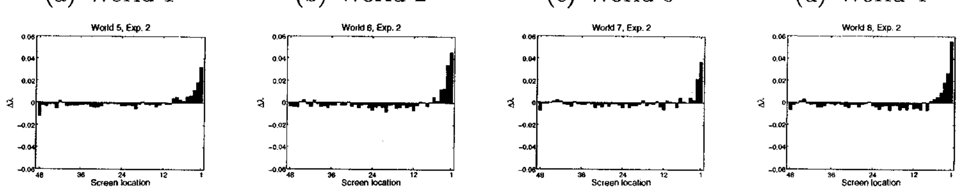


(a) World 1 (b) World 2

Figure 6.25: $\Delta\lambda_{k,w}$ for the two social influence worlds in experiment 2.



(a) World 1 (b) World 2 (c) World 3 (d) World 4



(e) World 5 (f) World 6 (g) World 7 (h) World 8

Figure 6.26: $\Delta\lambda_{k,w}$ for the eight social influence worlds in experiment 2.

6.4 Are demographics groups different in their behavior?

A natural question to explore is whether participants in different demographic groups differed in their behavior. To summarize our findings, there were not large differences across any of the demographic categories that we examined: gender, age, country of residence, self-reported opinion leader status, location during experiment, and type of Internet connection. However, there were some differences which we will now discuss.

Figure 6.27 plots the mean number of listens and downloads in all four experiments for men and women and shows that men listened to slightly more songs on average, and that women had a higher probability of downloading a song to which they were listening. Figure 6.28 plots these same results for each of the four experiments separately and shows that these patterns were similar across all four experiments.

Figure 6.29 plots the mean number of listens and downloads in all four experiments for subjects under 18, 18-34, and 35+ and shows that older participants listened to more songs on average and had a lower probability of downloading a song to which they were listening. Figure 6.30 plots these same results for each of the four experiments separately and shows that these patterns were especially pronounced in experiment 1.

Figure 6.31 plots the mean number of listens and downloads in all four experiments for subjects living in the United States, UK/Canada, and other countries and shows that subjects from other countries listened to more songs on average and had a lower probability of downloading a song to which they were listening. Figure 6.32 plots these same results for each of the four experiments separately and shows that these patterns were similar across experiments.

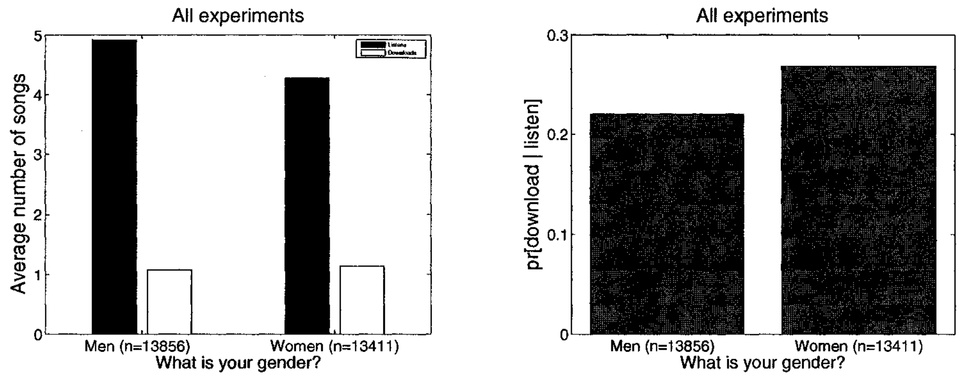
Figure 6.33 plots the mean number of listens and downloads in all four experiments by self-reported opinion leader status and shows that behavior was similar across groups.[8] Further, figure 6.34 plots these same results for each of the four experiments separately and shows again that these patterns were similar across experiments.

Figure 6.35 plots the mean number of listens and downloads in all four experiments for subjects by their type of Internet connection and shows, perhaps not surprisingly, that subjects using broadband connections listened to more songs on average. However, somewhat surprisingly, subjects who were using dial-up connections were more likely to download a song to which they were listening. Figure 6.36 plots these same results for each of the four experiments separately and shows that the largest differences between these two groups were observed in experiment 1.

Finally, figure 6.37 plots the mean number of listens and downloads for subjects in all four experiments by their location while participating (home, office, school, other) and shows subjects participating from the office listened to the most songs on average while subjects participating from school listened to the fewest. However, subjects listening from the office were the least likely to download a song to which they were listening. Figure 6.38 plots these same results for each of the four experiments separately and shows that these patterns were similar across experiments.

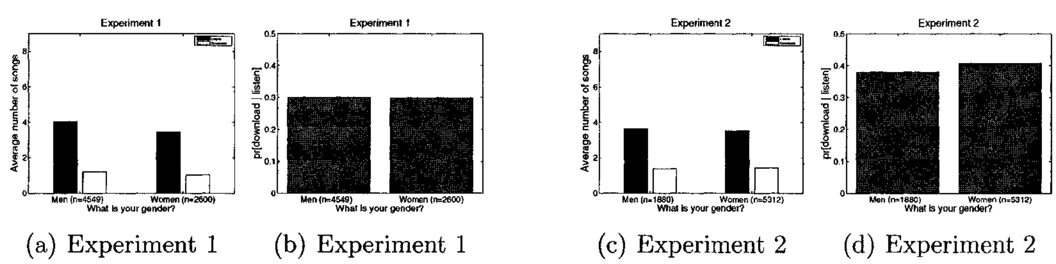
To conclude, we observed some small differences across the demographic groups, but nothing that created any questions for further follow-up or that would call into question any of the previous findings.

[8]Self-reported opinion leader status was, following Katz and Lazarsfeld (1955, Ap. B), assessed with the following question, "Compared to your circle of friends, how likely are you to be asked for advice about music? [much less likely], [less likely], [more likely], [much more likely]."



(a) (b)

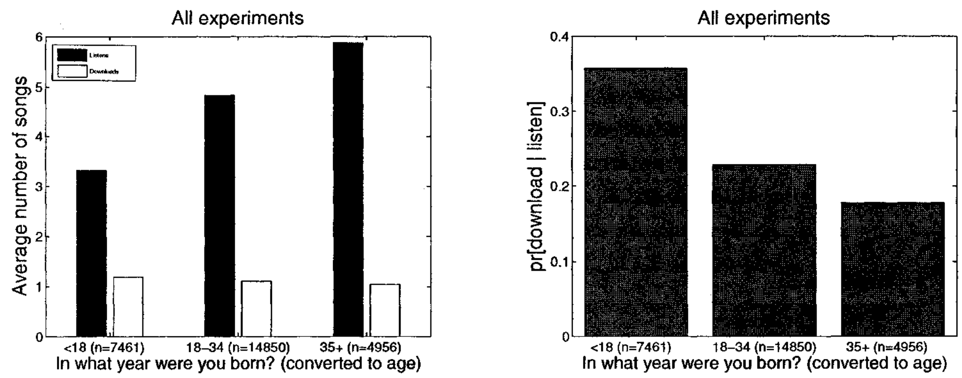
Figure 6.27: Average number of listens and downloads by gender and probability of download given listen by gender in all four experiment together.



(a) Experiment 1 (b) Experiment 1 (c) Experiment 2 (d) Experiment 2

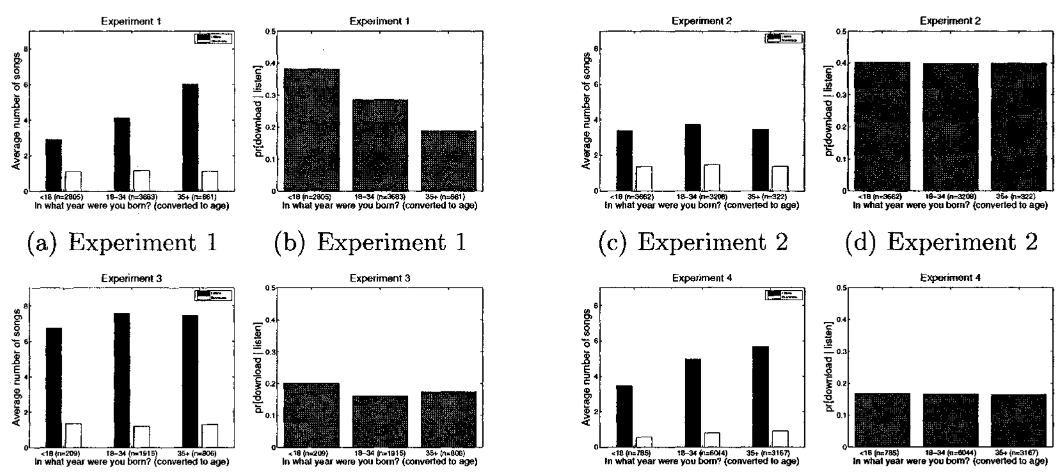
(e) Experiment 3 (f) Experiment 3 (g) Experiment 4 (h) Experiment 4

Figure 6.28: Average number of listens and downloads by gender and probability of download given listen by gender in each of the four experiment separately.



(a) (b)

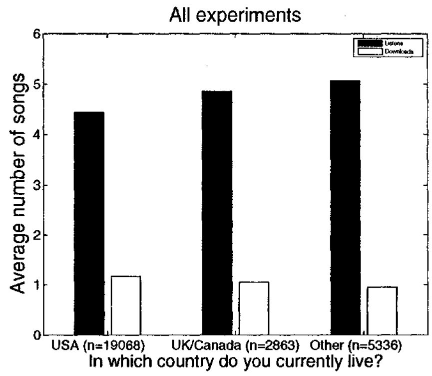
Figure 6.29: Average number of listens and downloads by age and probability of download given listen by age in all four experiment together.



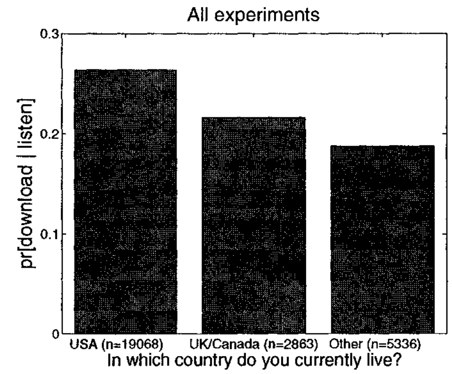
(a) Experiment 1 (b) Experiment 1 (c) Experiment 2 (d) Experiment 2

(e) Experiment 3 (f) Experiment 3 (g) Experiment 4 (h) Experiment 4

Figure 6.30: Average number of listens and downloads by age and probability of download given listen by age in each of the four experiment separately.

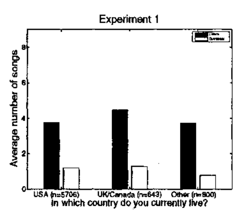


(a)

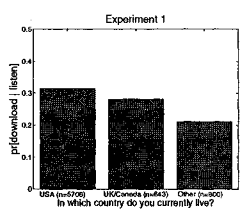


(b)

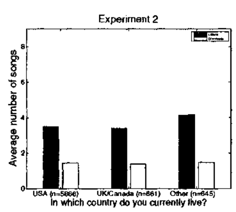
Figure 6.31: Average number of listens and downloads by country and probability of download given listen by country in all four experiment together.



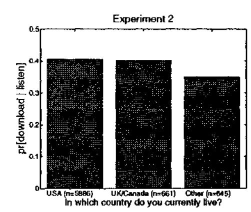
(a) Experiment 1



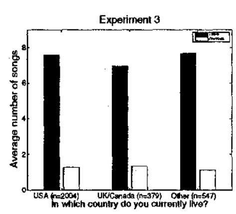
(b) Experiment 1



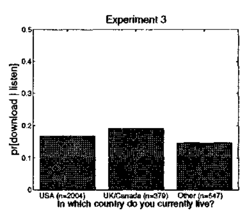
(c) Experiment 2



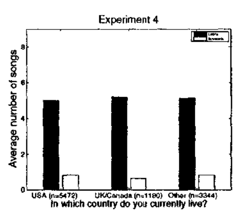
(d) Experiment 2



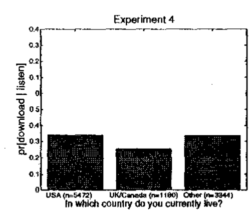
(e) Experiment 3



(f) Experiment 3

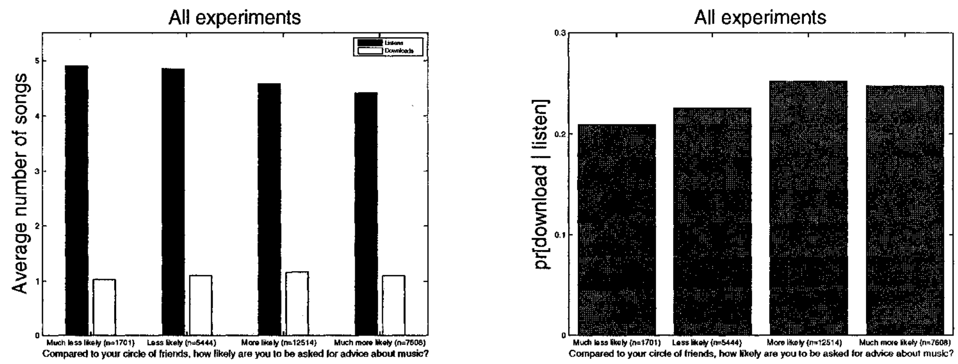


(g) Experiment 4



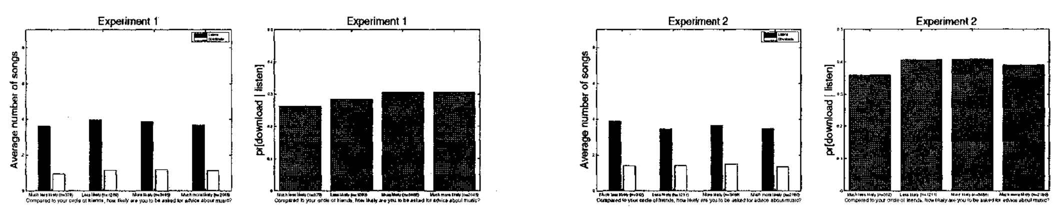
(h) Experiment 4

Figure 6.32: Average number of listens and downloads by country and probability of download given listen by country in each of the four experiment separately.

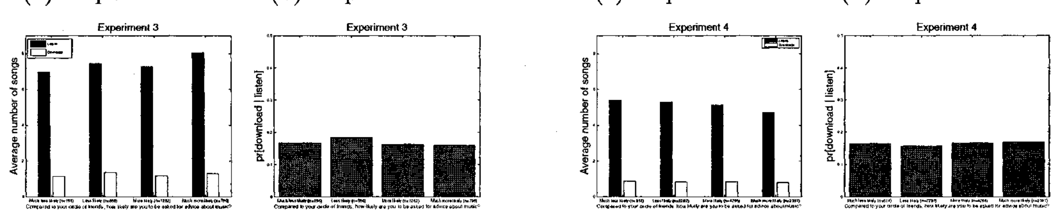


(a) (b)

Figure 6.33: Average number of listens and downloads by self-reported "opinion leadership" and probability of download given listen by self-reported "opinion leadership" in all four experiment together.

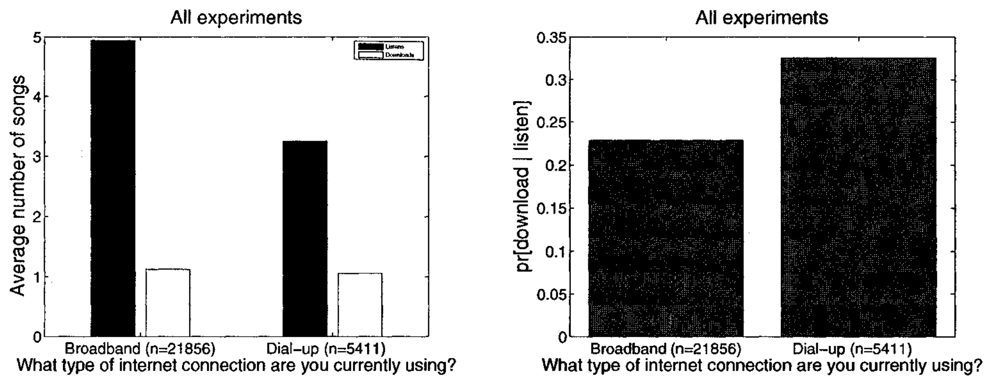


(a) Experiment 1 (b) Experiment 1 (c) Experiment 2 (d) Experiment 2



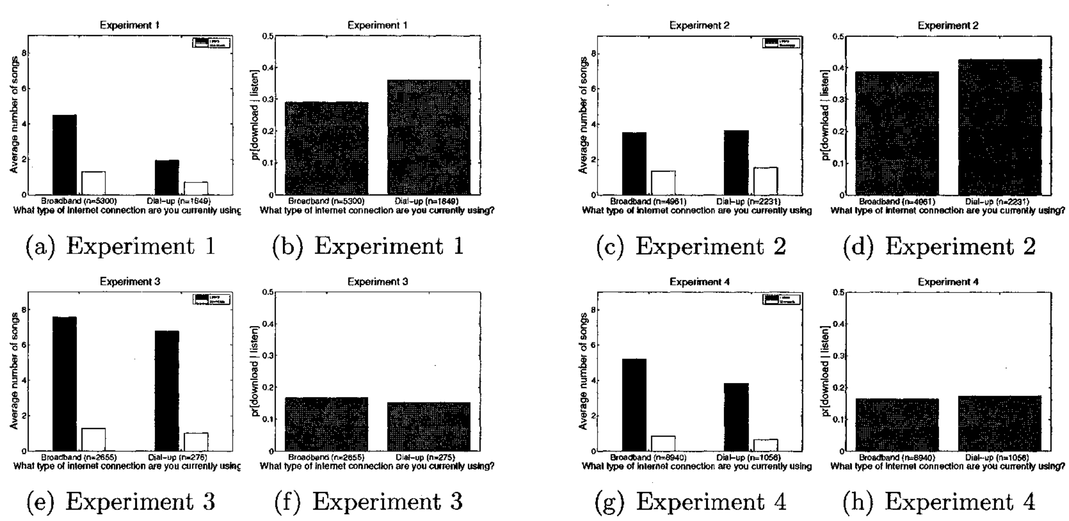
(e) Experiment 3 (f) Experiment 3 (g) Experiment 4 (h) Experiment 4

Figure 6.34: Average number of listens and downloads by self-reported "opinion leadership" and probability of download given listen by self-reported "opinion leadership" in each of the four experiment separately.



(a) (b)

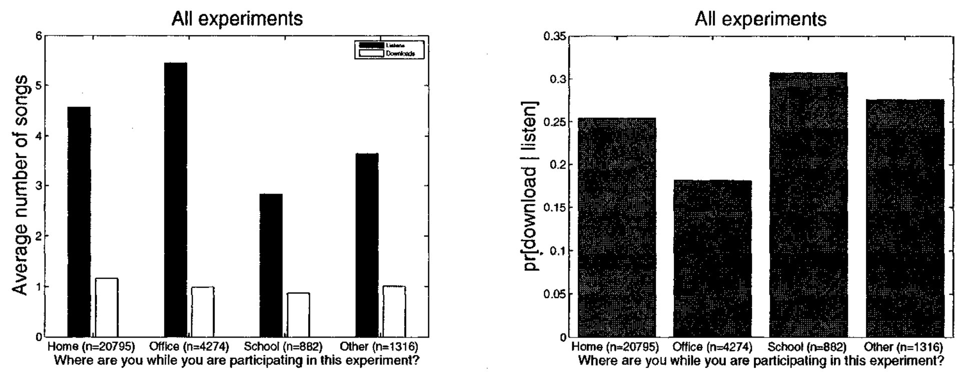
Figure 6.35: Average number of listens and downloads by connection and probability of download given listen by connection in all four experiment together.



(a) Experiment 1 (b) Experiment 1 (c) Experiment 2 (d) Experiment 2

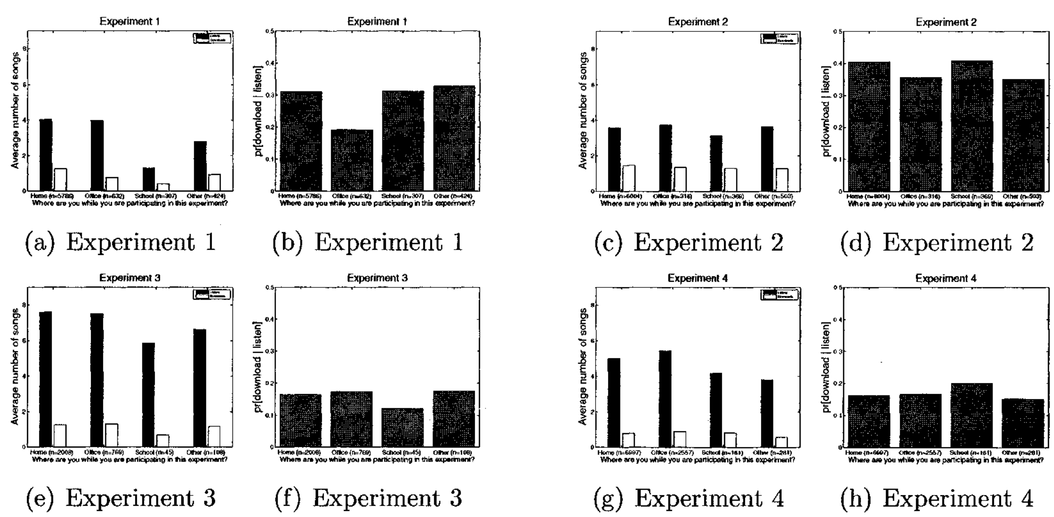
(e) Experiment 3 (f) Experiment 3 (g) Experiment 4 (h) Experiment 4

Figure 6.36: Average number of listens and downloads by connection and probability of download given listen by connection in each of the four experiment separately.



(a) (b)

Figure 6.37: Average number of listens and downloads by location and probability of download given listen by location in all four experiment together.



(a) Experiment 1 (b) Experiment 1 (c) Experiment 2 (d) Experiment 2

(e) Experiment 3 (f) Experiment 3 (g) Experiment 4 (h) Experiment 4

Figure 6.38: Average number of listens and downloads by location and probability of download given listen by location in each of the four experiment separately.

6.5 Are some demographic groups more susceptible to social influence?

Individuals are thought to vary in their susceptibility to social influence (Bearden et al., 1989; Pasupathi, 1999). Here we can attempt to quantify the social influence on listening decisions. We begin by writing the likelihood function for α, the parameter that will be used to measure social influence.

$$L(\alpha) = \prod_{i=1}^{n} p(\vec{l_i} \mid \vec{d_i}, \alpha) \tag{6.7}$$

where $\vec{l_i}$ is the vector of listen decisions for participant i and $\vec{d_i}$ is the vector of download counts that participant i sees. Following standard practice, we can take the log of both sides,

$$log(L(\alpha)) = \sum_{i=1}^{n} log(p(\vec{l_i} \mid \vec{d_i}, \alpha)) \tag{6.8}$$

To proceed we need a model of $p(\vec{l_i} \mid \vec{d_i}, \alpha)$. One natural model is to assume that the probability of any listen vector is simply the product of the individual listening choices which make up that vector,

$$p(\vec{l_i} \mid \vec{d_i}, \alpha) = \prod_{j=1}^{48} p(l_{ij}) \tag{6.9}$$

where

$$p(l_{ij}) = \begin{cases} \frac{(d_{ij}+\varepsilon)^\alpha}{\sum_{j=1}^{48}(d_{ij}+\varepsilon)^\alpha} & \text{if } \ l_{ij} = 1 \\ 1 - \frac{(d_{ij}+\varepsilon)^\alpha}{\sum_{j=1}^{48}(d_{ij}+\varepsilon)^\alpha} & \text{if } \ l_{ij} = 0 \end{cases}$$

where d_{ij} is the download count of song j that person i sees and ε is a small positive constant which is included so that there is a small probability of someone listening to a song with 0 downloads. Under this model if $\alpha = 0$ then all songs are equally likely to be listened to and as α increases the probability of listening to the more popular songs increases. For example, if $\alpha = 1$ the probability of listening to a songs is proportional to their number of downloads.

Although this seems a sensible model, preliminary exploration of equation 6.9 revealed an important problem: there was a strong relationship between the number of listens and the probability of a given listen vector, $\vec{l_i}$. For example, assume that $\alpha = 1$ (i.e. the probability of listening to a song is proportional to the number of previous downloads). Under this situation the probability of not listening to any given song is much greater than the probability of listening to that song.[9] Thus, vectors with a few listens had a much higher probability than vectors with many listens. Therefore, the likelihood was dominated by participants who had few listens. It did not seem fair to weight participants differently in estimating α so for estimation we only used participants' first listen which removes these problems.[10] Note, that this comes at a huge cost of throwing out tons of data, and therefore, is clearly not ideal. Given that caveat, however, we will proceed.

To build intuition we will first estimate α in each experiment.[11] Figure 6.39 plots the log-likelihood of α in the four different experiments. For experiment 1, $\alpha \approx 0.4$ is most consistent with the observed data (figure 6.39(a)). For experiment 2, where we would expect social influence to be stronger, we find indeed that the estimate increases to $\alpha \approx 1.5$ (figure 6.39(b)). For experiment 3, the estimate decreases so that $\alpha \approx 0.7$ is most consistent with the data (figure 6.39(c)). Finally, for experiment 4, we estimate $\alpha \approx 1$ (figure 6.39(d)).[12]

---

[9] That is, $\frac{(d_{ij}+\varepsilon)^\alpha}{\sum_{j=1}^{48}(d_{ij}+\varepsilon)^\alpha} \ll 1 - \frac{(d_{ij}+\varepsilon)^\alpha}{\sum_{j=1}^{48}(d_{ij}+\varepsilon)^\alpha}$.

[10] An alternative way to get around this problem is to somehow condition on the number of listens in the vector. For example let $\mathcal{S}_n$ be the set of all vectors, $\vec{l_i}$, with n listens. We would like to constrain things such that $\sum_{\vec{l_i} \in \mathcal{S}_n} p(\vec{l_i} \mid \vec{d_i}, \alpha) = 1$. However, for a modest number of listens this calculation is impossible. For example, $\mathcal{S}_{20}$ has $\binom{48}{20}$ elements (i.e. 1×10^{13}).

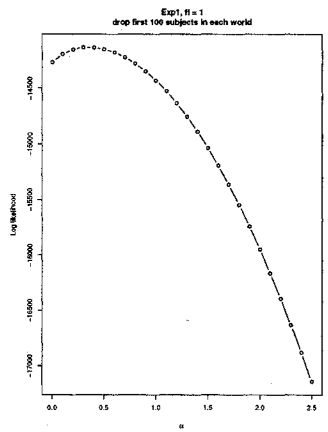
[11] Note that for all estiamtes we dropped the first 100 participants in each world.

[12] Since we are only analyzing the first listen it is unlikely that the inversion affected the results. If we analyzed all listens, however, one might expect different patterns in the inverted worlds as subjects realized that popularity was not strongly related to appeal.

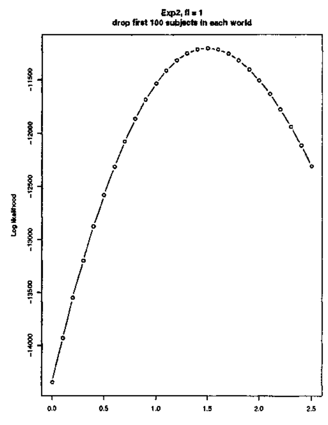
To make the interpretation of these values of α more clear imagine two songs, one with 100 downloads and one with 10. If $\alpha = 0$ a participant is equally likely to listen to both. If $\alpha = 0.4$, as in experiment 1, the participant is 2.5 times more likely to listen to the more popular song ($2.5 \approx \frac{100^{0.4}}{10^{0.4}}$). However, if $\alpha = 1.5$, as in experiment 2, the participant is about 30 times more likely to listen to the more popular song ($30 \approx \frac{100^{1.5}}{10^{1.5}}$). Thus, this detected difference in α reflects a substantial behavior change, one that was expected given our experiment design (recall that in experiment 1 the songs were presented in a grid unsorted by popularity and that in experiment 2 they were presented in one column sorted by popularity).

The above estimates are based on results from all social influence worlds together. Figures 6.40, 6.41, 6.42 and 6.43 plots the estimates from each world in each experiment separately. Except for experiment 1, these results show that the estimated α in the independent world is close to 0 and much smaller than the estimated α in the social influence worlds.[13] The estimated α in the social influence worlds also appear to be similar across worlds. Collectively, these results give us at least some confidence in the results from the model.

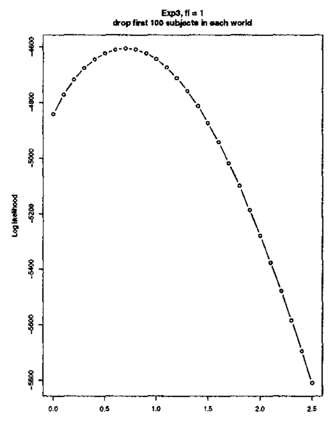
[13] We are not sure why experiment 1 would be different in this regard.



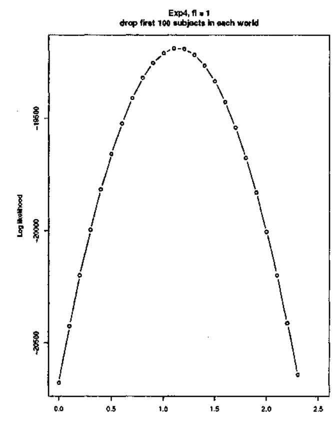
(a) Experiment 1



(b) Experiment 2



(c) Experiment 3



(d) Experiment 4

Figure 6.39: Estimated α in experiments 1, 2, 3, and 4.

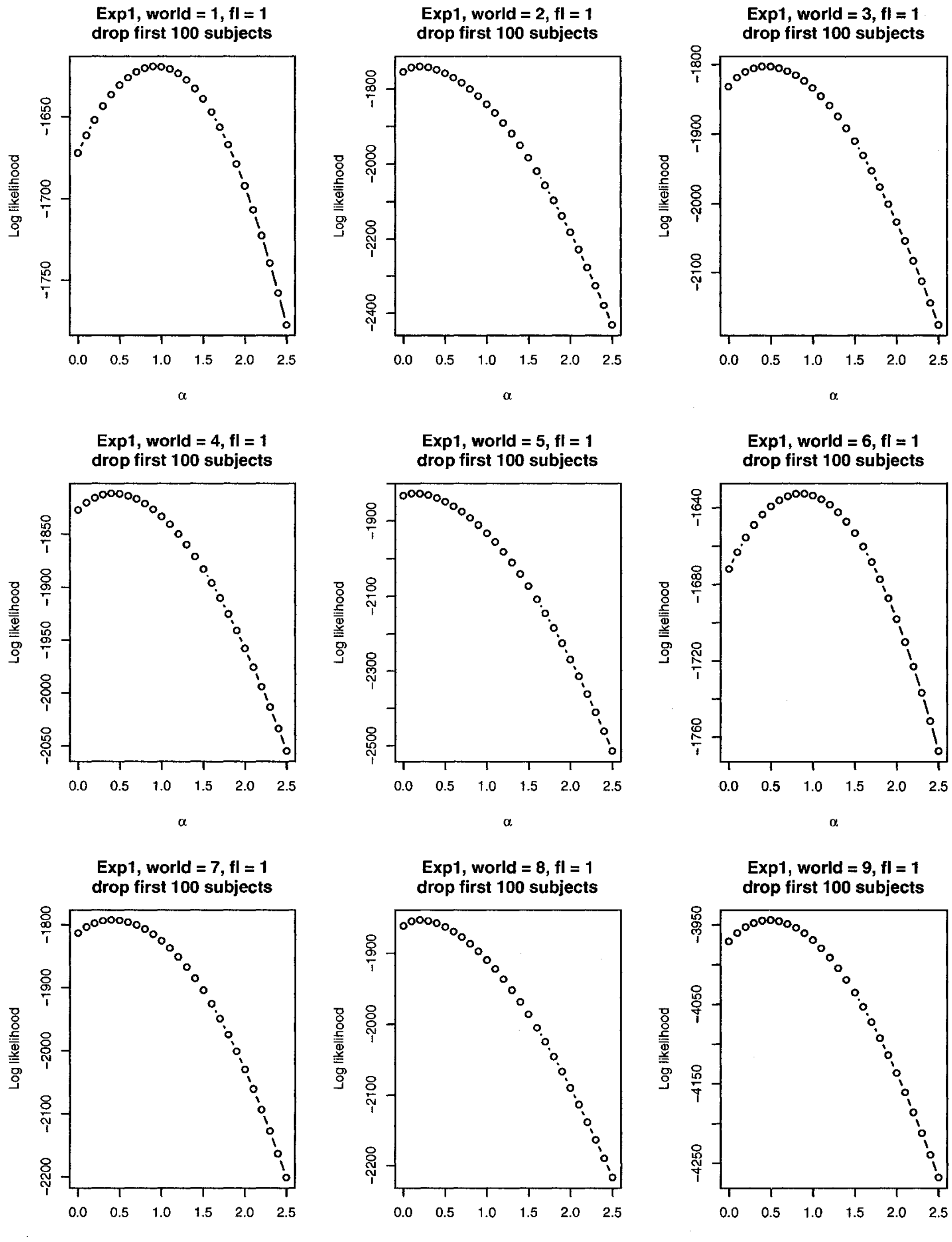


Figure 6.40: Estimated α in each world in experiment 1. We are not sure why the estimated α in the independent world (world = 9) is neither much lower than the estimates in the social influence worlds nor close to 0.

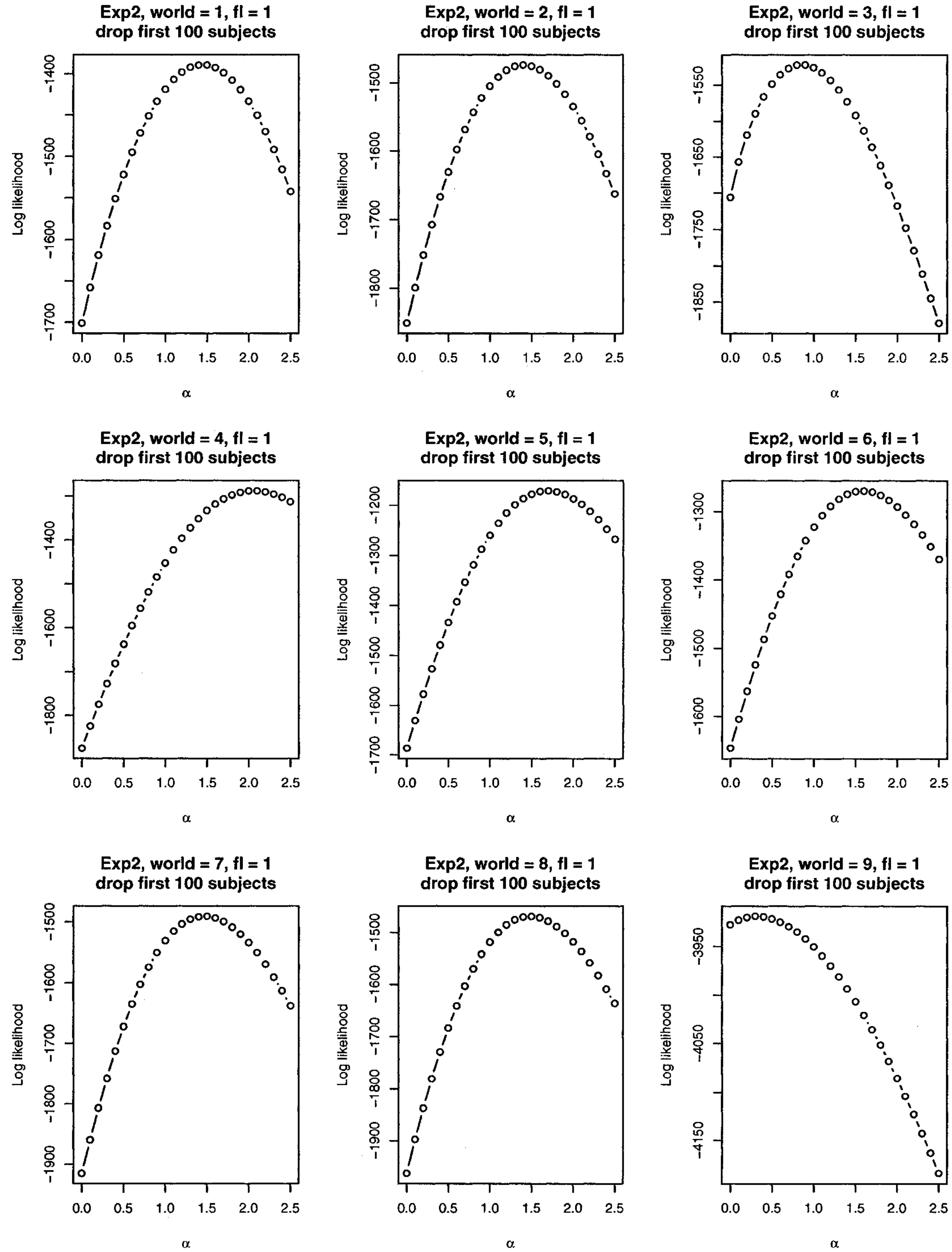


Figure 6.41: Estimated α in each world in experiment 2. As expected the estimated α in the independent world (world = 9) is lower than the estimates in the social influence worlds, but we are not sure why it is not closer to 0.

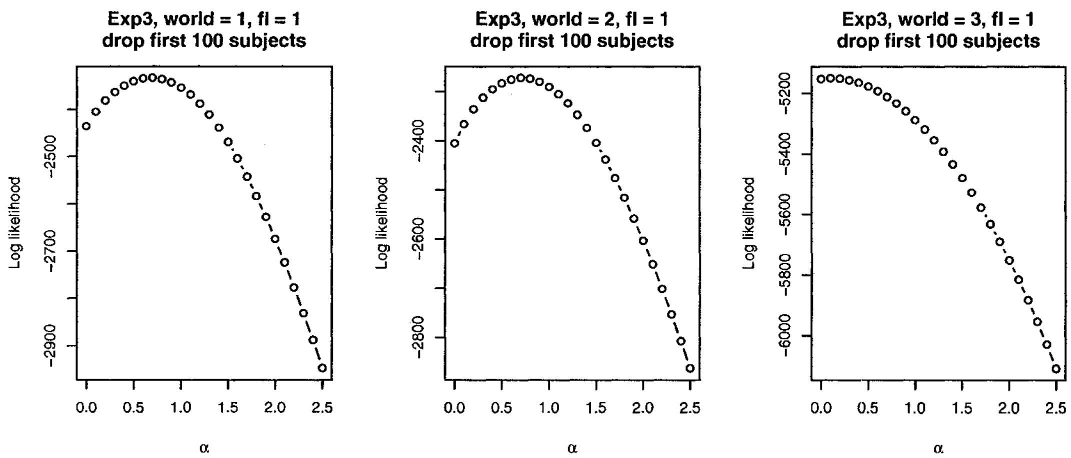


Figure 6.42: Estimated α in each world in experiment 3. As expected the estimated α in the independent world (world = 3) is close to 0.

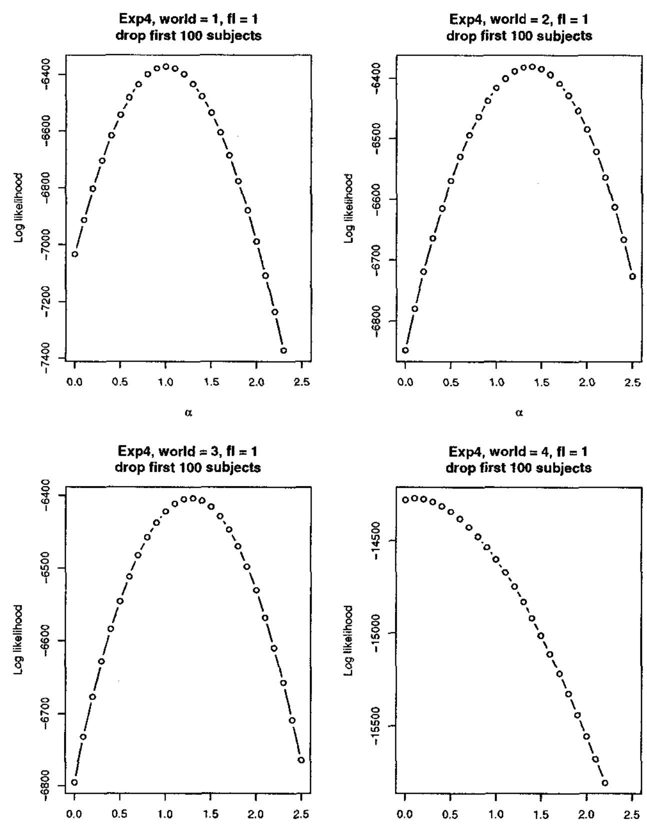


Figure 6.43: Estimated α in each world in experiment 4. As expected the estimated α in the independent world (world $= 4$) is close to 0.

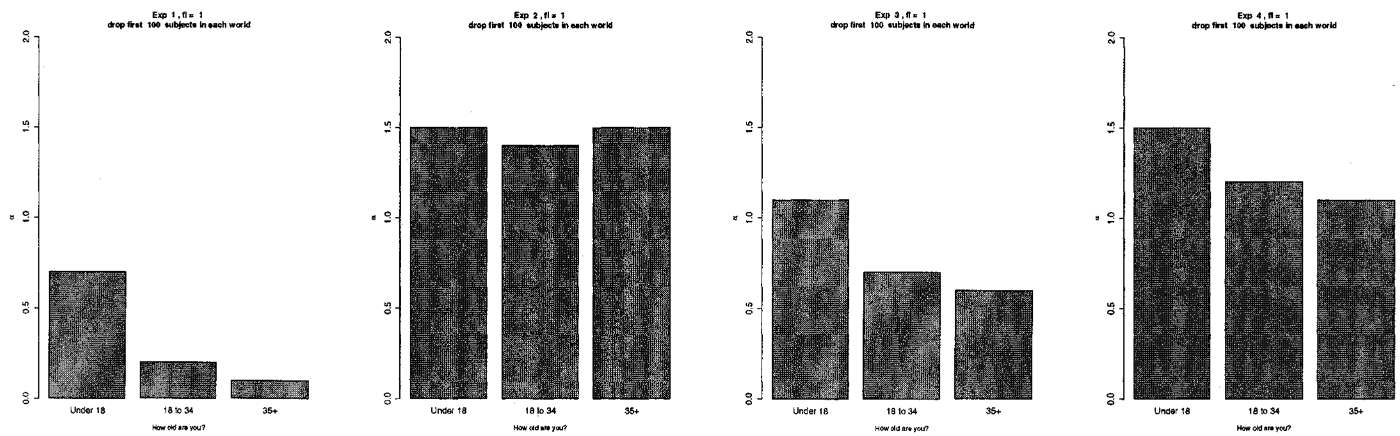
6.5.1 Estimating α for different subgroups

Given that the estimated values of α are reasonable when considering all subjects, we will turn our attention to estimating α for different subgroups in the population by refitting the model on different subsets of data. It would probably be possible to have a more sophisticated estimation approach that jointly estimates all parameters, but that was not undertaken here.

Generally we found that between-experiment variation in α is bigger than within-experiment variation across subgroups. Further, there don't seem to be consistent patterns across experiments, except for patterns between age groups. Figure 6.44 plots the estimated α for participants under 18, between 18 and 34, and 35 plus. In experiments 1, 3 and 4 there is monotonic pattern where younger participants seem to have a higher α indicating that their listening decisions were more influenced by popularity. This finding is consistent with previous experimental results from a different context (Pasupathi, 1999).

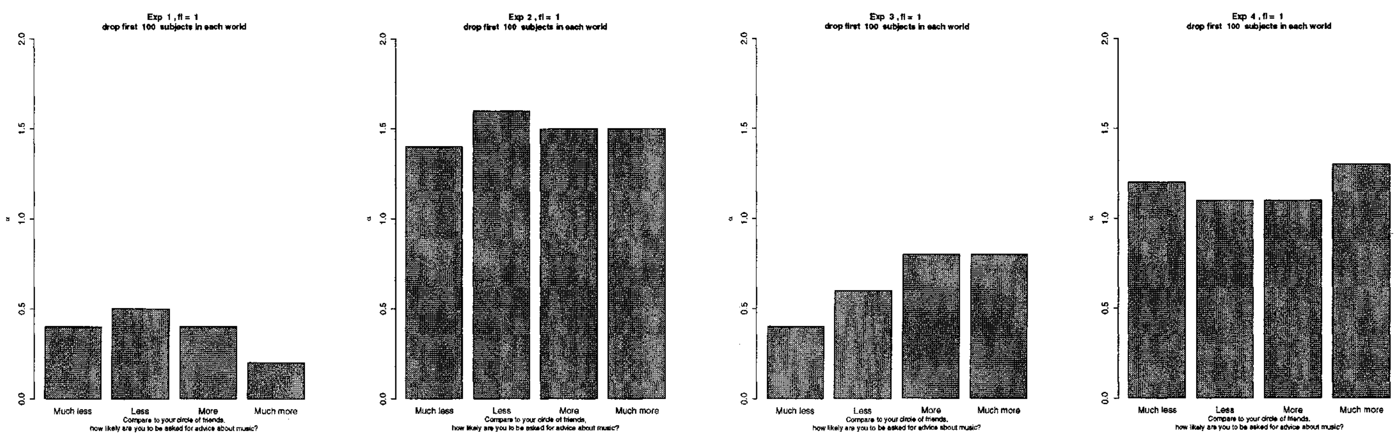
Across the other demographic groups that we examined, however, there were not consistent patterns. Figure 6.45 report α by self-reported opinion leader status.[14] We find no consistent pattern across experiments. Figure 6.46 reports the same estimates by gender and again does not find a consistent pattern; sometimes men had higher estimated α while sometimes women had high estimated α (even when there were differences, these differences were small). Figure 6.47 estimates α for the different participants based on their location while completing the experiment. Again we find no consistent pattern. The extremely high value for α that was estimated for participants at school in experiment 3 should be ignored because it is only based on 10 participants. Finally, figure 6.48 shows that we did not find consistent differences between participants in different countries.

[14]Self-reported opinion leader status was, following Katz and Lazarsfeld (1955, Ap. B), assessed with the following question, "Compared to your circle of friends, how likely are you to be asked for advice about music? [much less likely], [less likely], [more likely], [much more likely]."



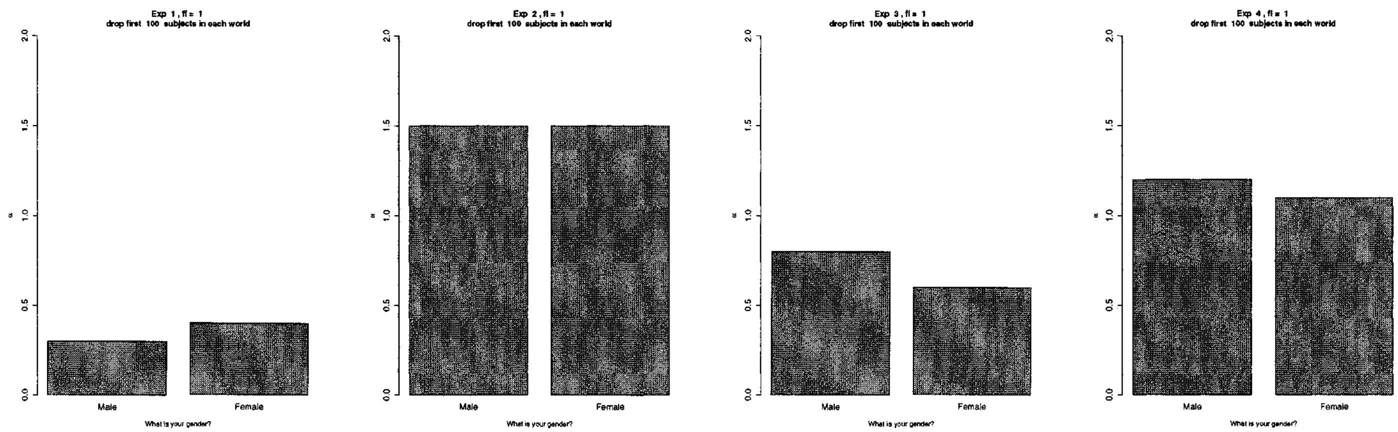
(a) Experiment 1 (b) Experiment 2 (c) Experiment 3 (d) Experiment 4

Figure 6.44: Estimated α by age in experiments 1, 2, 3, and 4.



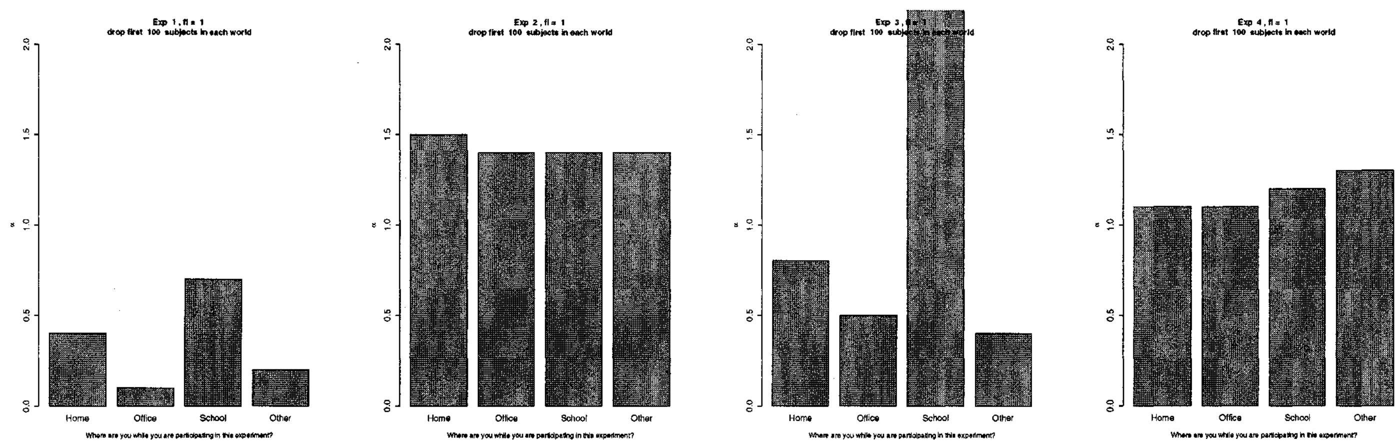
(a) Experiment 1 (b) Experiment 2 (c) Experiment 3 (d) Experiment 4

Figure 6.45: Estimated α by self-reported opinion leader status in experiments 1, 2, 3, and 4.



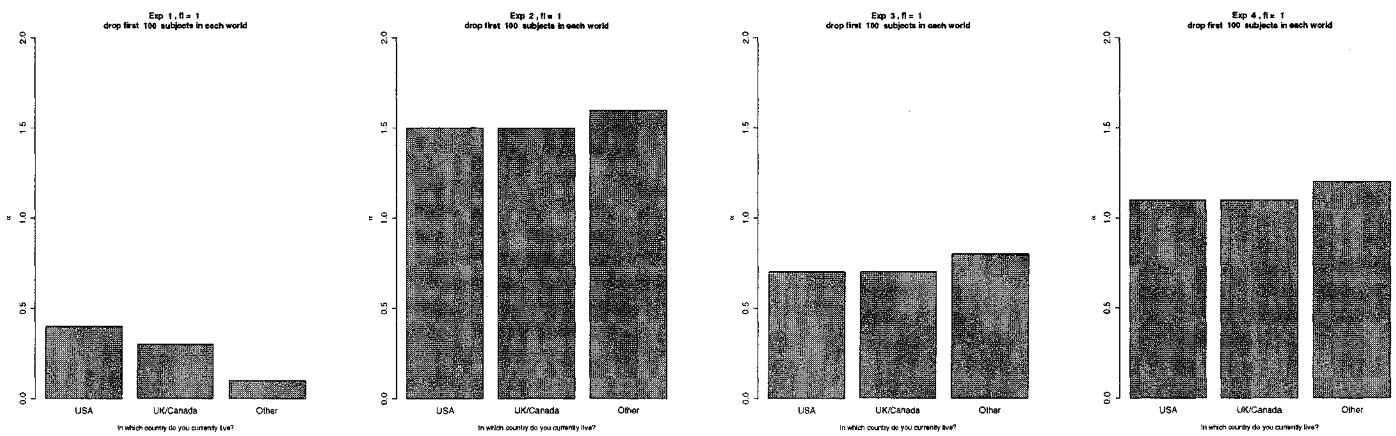
(a) Experiment 1 (b) Experiment 2 (c) Experiment 3 (d) Experiment 4

Figure 6.46: Estimated α by gender in experiments 1, 2, 3, and 4.



(a) Experiment 1 (b) Experiment 2 (c) Experiment 3 (d) Experiment 4

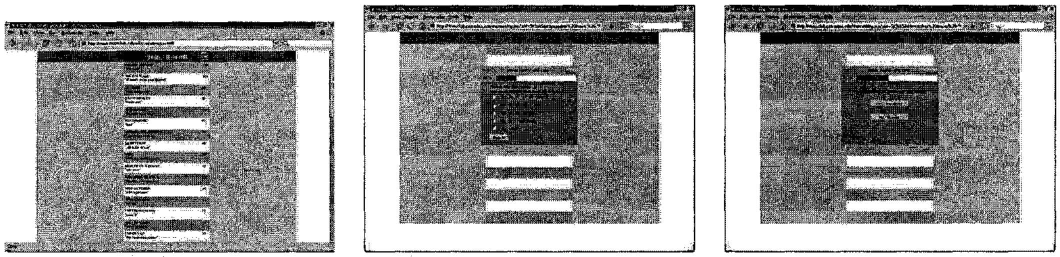
Figure 6.47: Estimated α by location in experiments 1, 2, 3, and 4. The extremely high value for individuals at school observed during experiment 3 should be ignored because it is only based on 10 people.



(a) Experiment 1 (b) Experiment 2 (c) Experiment 3 (d) Experiment 4

Figure 6.48: Estimated α by country in experiments 1, 2, 3, and 4.

To conclude, this analysis was extremely crude, and therefore, extreme care should be taken in interpreting the results. First, this analysis was based only on participants' first listen (for reasons described previously) and so does not use a large portion of the available data. Further, no checks of model fit have been performed, even though these are very important when estimating statistical models (Gelman et al., 2004; Gelman and Hill, 2007). Finally, no standard errors are provided making interpretation of the parameters difficult. Despite these huge limitations, however, we still believe that the non-findings here are meaningful in a practical sense. Even thought our analysis is blunt, it was able to detect large effects, for example, the difference between experiment 1 and 2 (figure 6.39). Thus, except for the age groups, the lack of meaningful differences between subgroups in terms of the susceptibility to social influence when deciding which song to listen to suggests that these differences are probably small, at least in the context of this experiment.



(a) Listen-choice screen (b) Rating-choice screen (c) Download-choice screen

Figure 6.49: Screenshots of the listen-choice, rating-choice, and download-choice screen. Popularity is only shown on the listen-choice screen.

6.6 Was there social influence on the rating and download decision?

We have shown many places in this dissertation that subjects' listening decisions were influenced by the popularity of the songs. But were subjects rating and download decisions also influenced? To summarize the results that will follow, we did not find strong evidence of social influence on the rating and download decisions, but our experiment was not designed to detect this influence. We believe that we did not find this influence because, given the nature of our website, the popularity of the song was not salient when participants were making the rating and download decisions (for screenshots see figure 6.49).

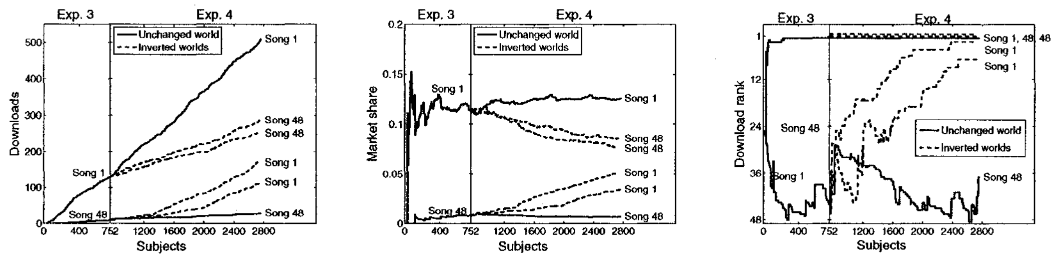
As mentioned briefly before, it is important to note that our experimental design is not ideal to measure social influence on the rating and download decisions because in our experiments, subject chose which song to listen to and therefore the pool of people making rating and download decisions about the songs may differ. To further understand this point, we can compare our design to the between-subjects design of Huang and Chen (2006) which was meant to explore the effect of popularity on consumer choices. In the experiment of Huang and Chen, consumers had to indicate their likelihood of purchasing one of two travel books—"Happy Travel" and

"Easy Travel"—in one of three different conditions. In one condition subjects were told that "Happy Travel" was much more popular, in another subjects were told that "Happy Travel" was slightly more popular, and in the final condition subjects were told that the books were equally popular. Consistent with expectations, subjects reported a higher likelihood of purchasing "Happy Travel" in the conditions where it was more popular. Because subjects were randomly assigned to condition, Huang and Chen were able to isolate this popularity effect. In our experiments, however, subjects chose which songs to listen to, and therefore which songs they would rate and download. Because different types of people may chose to listen to popular or unpopular songs, the pool of participants making ratings and download decisions about each song may vary. These differences ultimately prevent us from precisely isolating social influence on the rating and download decision.[15]

Given this caveat, we will proceed. Here we will look for social influence on rating and download decisions where it is most likely to occur, in situations where the popularity of a given song varied dramatically across different worlds. If the song gets higher ratings or has a higher batting average when it is more popular, that indicates that social influence may be operating on these decisions.[16] This difference in popularity across worlds occurred as a result of our inversion (recall Chapter 5). To review briefly, in the unchanged world song 1 ("She Said" by Parker Theory) was the most popular song, but in the inverted worlds, it was initially presented as the least popular song, although as the experiment progressed its popularity increased. This can be compared to song 48 ("Florence" by Post Break Tragedy) which was

[15]This inability to isolate social influence on the rating and download decision is not a flaw in our experimental design. Rather, recall that our experiments were conducted to explore the role of individual level social influence on collective outcomes, and our design was suited for that goal. Here we are attempting to use the data for a purpose that was not intended in the original design. If we really care about getting good answers about social influence on rating and download decisions, we would need to conduct separate, individual level, psychology-type experiments.

[16]Recall that any differences could also be because different people choose to listen to a song in the world where it is popular than the world where it isn't.



(a) Downloads (b) Market share (c) Download rank

Figure 6.50: Success dynamics of song 1 ("She Said" by Parker Theory) and song 48 ("Florence" by Post Break Tragedy) in experiment 4.

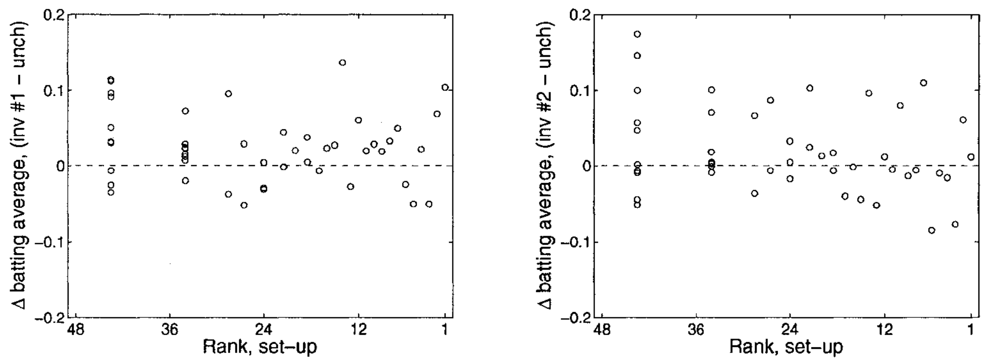
one of the least popular songs in the unchanged world, but the most popular song in the inverted worlds. The dynamics of popularity of these two songs are presented in figure 6.50.

First, as might be expected, song 48 had a higher mean rating and batting average in the inverted worlds than the unchanged world.[17] This suggests that when people think the song is popular, the song does better (i.e., higher mean rating and batting average). However, song 1 also had a higher mean rating and batting average in the inverted worlds, even thought it was less popular in these worlds.[18] This finding is contrary to what might be predicted by a simple social influence argument.

To understand how common it was for songs to do better in the inverted worlds, figure 6.51 presents the difference in batting average between the unchanged and inverted worlds as a function of the rank during the set-up period (i.e., positive values mean the song did better in the inverted world). This figures shows that almost all songs had higher batting averages in the inverted worlds, even those that

[17]The mean rating was 2.21 in the unchanged world compared to 2.43 and 2.38 in the inverted worlds (2.22 in the independent world); see also the dashed lines in figure E.11. The batting average was 0.103 in the unchanged world compared to 0.133 and 0.104 in the inverted worlds (0.089 in the independent world); see also the dashed lines in figure E.11

[18]The mean rating was 3.21 in the unchanged world compared to 3.40 and 3.43 in the inverted worlds (3.24 in the independent world); see also the dashed lines in figure 6.58. The batting average was 0.297 in the unchanged world compared to 0.393 and 0.364 in the inverted worlds (0.325 in the independent world); see also the dashed lines in figure 6.58



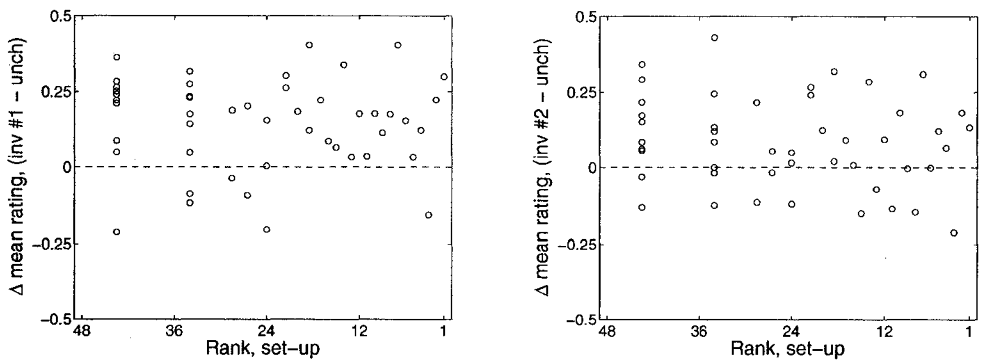
(a) Inv #1 - Unch (b) Inv #2 - Unch

Figure 6.51: Difference in batting average between inverted and unchanged as a function of the rank during the set-up period. Almost all songs had high batting average in the inverted worlds, even those that were made less popular by the inversion.

were hurt by the inversion. Figure 6.55 shows basically the same results for mean ratings; namely, that almost all songs did better in the inverted worlds. Another way to see this pattern whereby most songs doing better in the inverted worlds is with the multiple-thermometer plots presented in 6.53 and 6.54. These plots don't allow the direct comparison of any specific song, but they do show a general pattern that the distributions seem to be shifted in the inverted worlds.

However, a huge point of caution is required because it is not the case that the overall batting average and mean rating were higher in the inverted worlds. Instead, the opposite is true. The overall batting average in the unchanged world is 0.201, but only 0.176 and 0.171 in the inverted worlds. The overall mean rating is 2.71 in the unchanged world, but only 2.64 and 2.60 in the inverted worlds. The reason for this Simpson's paradox-type behavior is that in the inverted worlds, the lower appeal songs had more listens and so weighed more heavily on the overall batting average and overall mean rating.

Given the difference between the inverted and unchanged worlds, we can see if this was because people seemed to "like" the songs more in the inverted worlds



(a) Inv #1 - Unch (b) Inv #2 - Unch

Figure 6.52: Difference in mean rating between inverted and unchanged as a function of the rank during the set-up period. Almost all songs had higher mean rating in the inverted worlds, even those that were made less popular by the inversion.

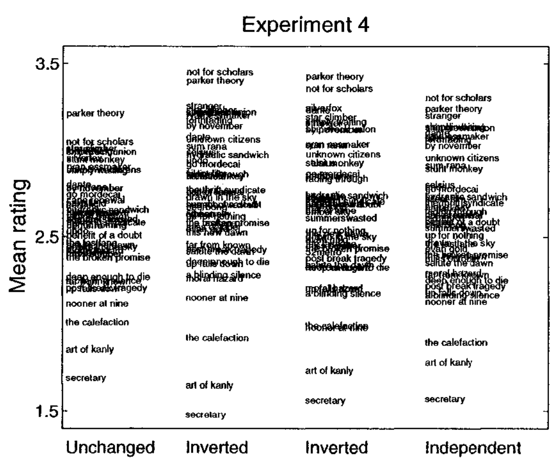


Figure 6.53: Multi-thermometer plot of mean rating for each song in each world in experiment 4.

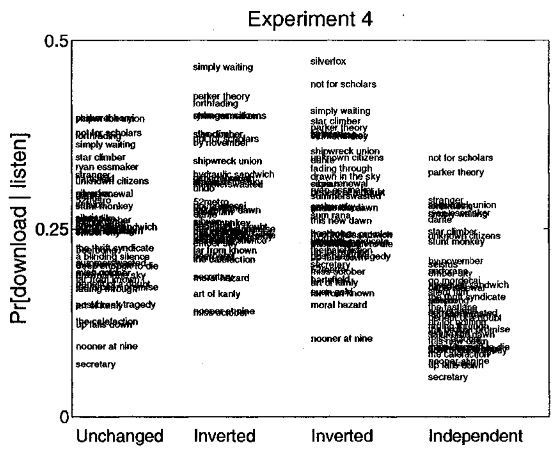
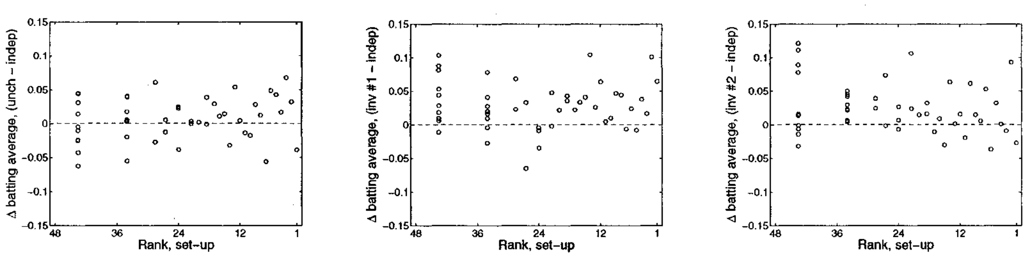


Figure 6.54: Multi-thermometer plot of batting average for each song in each world in experiment 4.

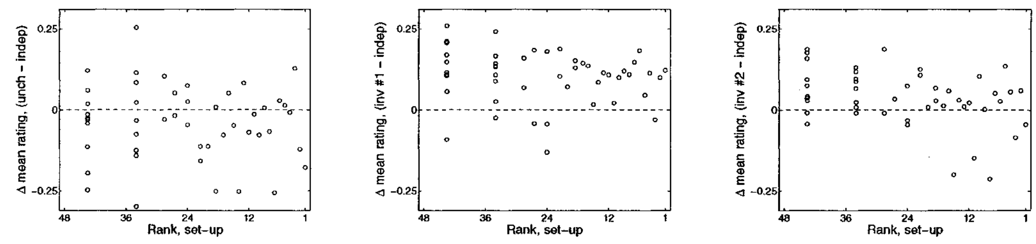


(a) Unch - Indep (b) Inv #1 - Indep (c) Inv #2 - Indep

Figure 6.55: Delta batting average between social influence worlds and independent world.

or if it was because people seemed to "like" the songs less in the unchanged world. To explore this we can compare the batting average and mean rating for all songs in the unchanged and inverted worlds to the results from the independent world, the result that represents our best measure of the "true" characteristics of the songs. Figure 6.55 shows that most songs do better in the inverted worlds, but that in the unchanged worlds, the batting averages are not systematically better or worse than the independent world. Figure 6.56 shows basically the same results for the ratings.

To further explore why the participants seem to "like" the songs more in the inverted world, we explored how these measures of "liking" changed over time in the



(a) Unch - Indep (b) Inv #1 - Indep (c) Inv #2 - Indep

Figure 6.56: Delta mean rating between social influence worlds and independent world.

four different worlds for song 1, a song that's popularity differed across the worlds and also changed dramatically in the inverted worlds. Figure 6.57 plots the moving average of the batting average with a window size of 500.[19] The dashed line represented the overall batting average[20] and the points on the y-axis represent the batting averages of all other songs in that world and thus allow the fluctuations in batting average to be compared to the variation within the population. For example, figure 6.57(a) shows that in the unchanged world, song 1 had the highest batting average and that this batting average was relatively constant. Figures 6.57(b) and 6.57(c) show that in

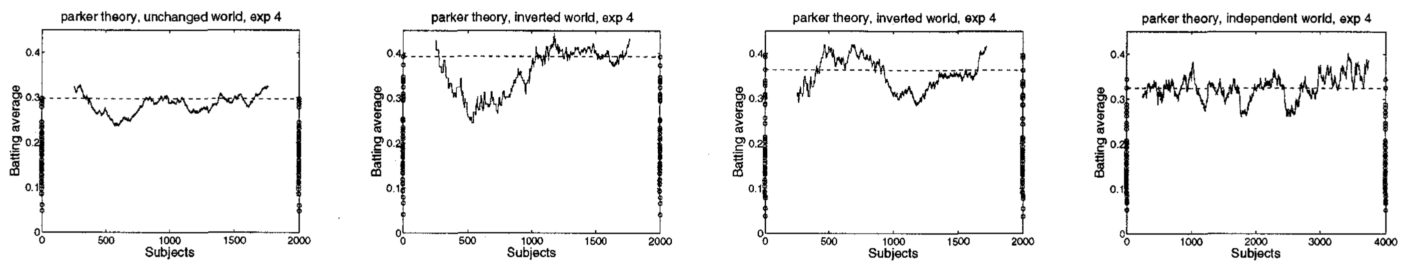
[19]So, for example, the first point on each graph ($x = 250$) represents the batting average over subjects 1 to 500. The second point ($x = 251$) represents the batting average of subjects 2 to 501, and so on. An alternative approach to the simple sliding window filter is to attempt to estimate an "instantaneous batting average" using a Hidden Markov Model (HMM) as in Aizen et al. (2004). To review, we have the download decisions of each person who listened to each particular song which can be stored as a vector $d_j = (d_1, d_2, \ldots d_{L_j})$ where L_j is the number of listens for song j. The goal of the HMM is then to find a vector of biases $b_j = (b_1, b_2, \ldots b_{L_j})$ that were most likely to produce the download sequence d_j. The techniques required to estimate b_j are described in Aizen et al. (2004) and Felzenszwalb et al. (2004). These authors opted for the HMM over the more simple sliding window filter because they wanted to be able to detect instantaneous changes in b_j. Because we have no reason to suspect that the values of b_j will change so quickly in our context, we opted for the simpler sliding window filter.

[20]A careful reader might note that that dashed line does not seem to bisect the moving average in a symmetric way. For example, in figure 6.57(b) it seems that it would be impossible for the moving average to be lower than overall average for most of the time. The reason for this pattern is that there are not an equal number of listens in each window so the overall average does not weight all time points equally. Returning to figure 6.57(b), this song had fewer listens earlier in the experiment in the inverted worlds because it was ranked near the bottom and so these early points figure less heavily in the overall average.

the inverted worlds, song 1 also had one of the highest batting averages, and that the batting average did not change dramatically, even thought the song gained popularity over this time (figure 6.50). Figure 6.58 presents results for the dynamics of the mean rating which are basically the same. Thus, even as the song became more popular, its batting average and mean rating did not change in any substantial way. Figure E.11 and E.10 plot these same results for song 48 which again show no substantial change over time.

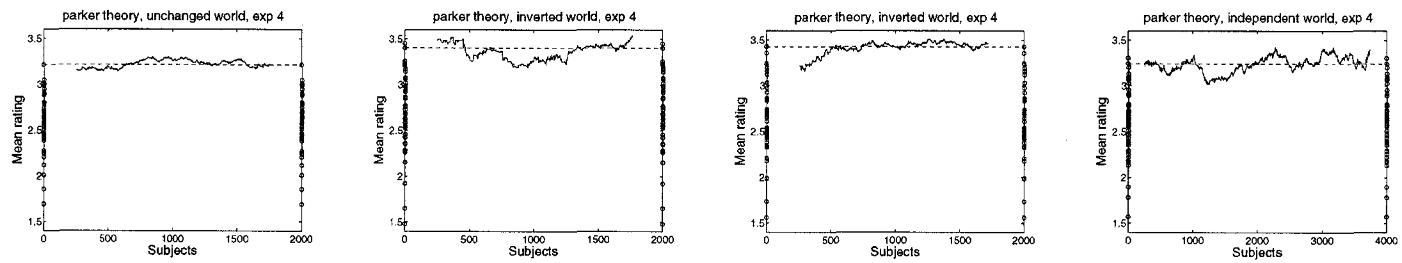
Returning to the original question, we do not think these data indicate strong social influence on the rating and download decision for two reasons. First, almost all songs did better in the inverted worlds, not just those that were helped by the inversion.[21] Second, for song 1, which started near the bottom of the inverted world and then gained popularity as the experiment progressed, we do not see an increase in either average rating or batting average. We think the there was little social influence on the rating and download decisions because the popularity of songs was not salient while participants where making these decisions (see screenshots in figure 6.49). All of these results should be interpreted with extreme caution because, as stated previously, our experimental design was suited for this analysis. In the end, we do think social influence probably affects how much people "like" cultural products in real markets and this possibility should be explored in additional psychology-type experiments.

[21]Ultimately it is not clear why many songs had a higher mean rating and batting average in the inverted worlds than the unchanged world, but we suspect that the participants listening to the songs in the different worlds were probably different.



(a) Unchanged (b) Inverted #1 (c) Inverted #2 (d) Independent

Figure 6.57: Dynamics of batting average for song 1 ("She Said" by Parker Theory) in the 4 worlds in experiment 4. This song had a higher batting average in the inverted worlds than the unchanged world. As time passed and the popularity improved in the inverted worlds, the batting average did not increase in a substantial way.



(a) Unchanged (b) Inverted #1 (c) Inverted #2 (d) Independent

Figure 6.58: Dynamics of mean rating for song 1 ("She Said" by Parker Theory) in the 4 worlds in experiment 4. This song had a higher mean rating in the inverted worlds than the unchanged world. As time passed and the popularity improved in the inverted worlds, the mean rating did not increase in a substantial way.

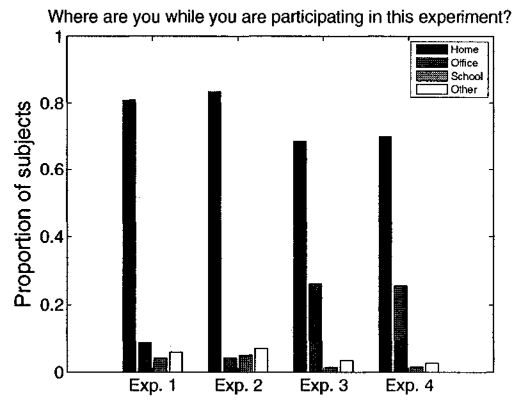


Figure 6.59: Subject location during the experiments.

6.7 Did these experiments tap into the "bored at work network"?

One major difference between web-based, as opposed to laboratory, experiments is subject recruitment. When planning this experiment we were hoping there would be word-of-mouth recruitment, perhaps tapping into what Jonah Peretti has called the "bored at work network," the huge pool of office workers sitting in cubicles and looking for something to do. Figure 6.59 plots the distribution of where participants reported being during our experiments. In experiments 1 and 2, less than 10 percent of participants reporting taking part in the experiment while at work; this figure increased to about 25 percent for experiments 3 and 4. Thus, it appears we did partially tap into the "bored at work network," but the great majority of participants where at home while participating. It is also relevant to point out again that people at work listened to and downloaded songs at a similar rate to those in other locations (figure 6.37).

On a related note, we also explored what days of the week were busiest for the experiment. Figure 6.60 plots the overall number of new registrants per weekday for the four experiments and shows the traffic was not substantially heavier during the

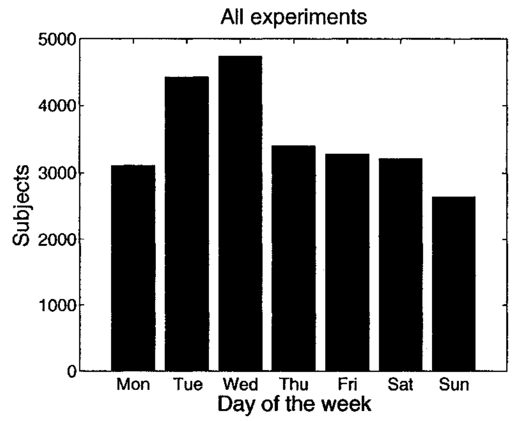
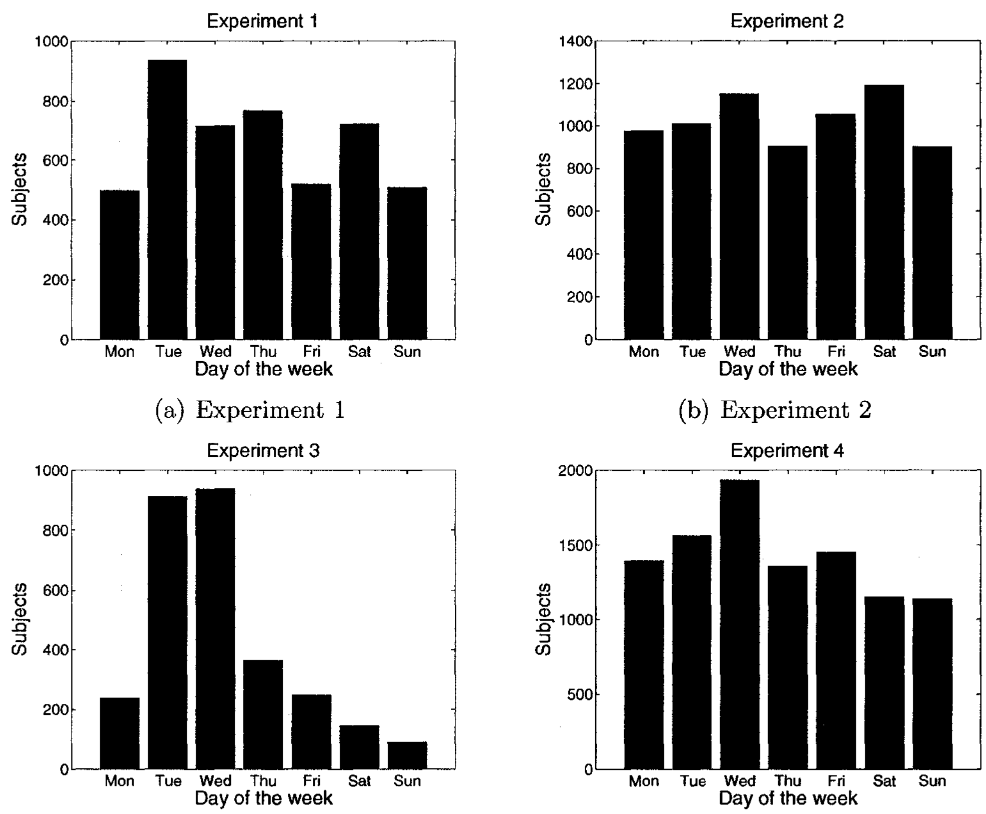


Figure 6.60: Subjects per day of the week in all four experiments. There doesn't seem to be much stronger traffic during the week.

weekdays. Figure 6.61 shows the histograms for each experiment separately. There does not seem to be large, systematic differences between the experiments. One potential anomaly was the spike in new registrations on Tuesday and Wednesday in experiment 3, but this was caused by our site getting mentioned on `www.boingboing.net` on Tuesday, April 5, 2005. Also, the slightly busier activity during the workweek in experiment 4 is probably due to our tendency to send out recruitment emails during the workweek.



(a) Experiment 1

(b) Experiment 2

(c) Experiment 3

(d) Experiment 4

Figure 6.61: Subjects per day of the week in each of the four experiments. There don't seem to be large systematic differences between the experiments. Note that the y-scales on the subfigures are different.

6.8 What are the implications of these findings for market design?

In addition to addressing a number of academic questions, we also believe that these experimental results, particularly experiment 4, have implications for the design of cultural markets. It is important to remember that the institutions that make up cultural markets are social constructions and there is no reason that they could not be constructed in some different way. For example, the "top 40" radio format was invented (some might say discovered) by a person at a specific time, namely, Todd Storz, a 1950's radio executive in Omaha, Nebraska (Fisher, 2007, p.6-7). Just as the top 40 format has become a stable institution, it is possible that there are other useful institutions for cultural markets that have yet to be developed, specifically in the new area of online cultural markets. We will now take a moment to describe a few ideas in this area.

Most online cultural markets face the problem of choice overload; for example, in 2003 it was estimated that `www.amazon.com` stocked 2.3 million book and 250,000 CDs (Brynjolfsson et al., 2003). These numbers have certainly increased and are probably dwarfed by the number of videos available on `www.youtube.com`. In order to help users filter through these many choices, websites often provide some information about the behavior of other users, just as we did in our experiments. Based on our results we can speculate about the effect of this information on consumer behavior and market dynamics.

First, it is useful to realize that there are several possible pieces of information that could be provided. From our experiment some possibilities included: number of listens, average rating, number of downloads, and batting average. Further, we could have provided any one of these results over a limited time window.[22] We have

[22]For example, best-seller lists and box-office grosses report all activity in a given week, rather

shown that providing users with the number of previous downloads increases the inequality and unpredictability of success (where success in measured by download counts). Further, we have shown that this feedback is susceptible to manipulation and not completely self-correcting.

From the perspective of a market designer the ideal social signal is one that 1) maximizes throughput (in our case downloads) and 2) is self-correcting to manipulation.[23] We believe that the best feedback to meet these goals would be either the average rating or the batting average (Aizen et al., 2004).

The findings from experiment 4 suggest that market throughput is positively related to the correlation between the signal and appeal; in other words, if the signal gets people to investigate things that they will enjoy. The average rating and the batting average both seem likely to push people to objects that they are more likely to enjoy. Further, both are self-correcting because if either measure is artificially inflated for a particular object, that song will receive more attention, but not more downloads or higher ratings, and thus the average rating or batting average will return to its previous level. Also, the larger the manipulation, the more attention and the quicker it will return to its previous value.

An alternative feedback method which might also work well would be to have some users randomly assigned into an independent condition and then have the behavior of these independent-condition users, shown to the other users. If the number of users is very large only a small fraction need be in the independent condition to provide useful information to the others.

We would also like to point out that some feedback measures currently in use, most notably the amount of previous attention (i.e., number of listens or views), are poor because they are not self-correcting. Namely, if someone manipulates the

than cumulative over the lifetime of the objects as was done in our experiment. These lists would not be particularly useful if they only provided lifetime figures.

[23]Note that from the perspective of an artist, other features might be desirable.

perceived amount of attention for an object, that will drive more people to that object, which will in turn drive more people, and so on. These popularity snowballs can also arise endogenously (without any manipulation) and will likely decrease the relationship between appeal and success. We believe that we would have observed even more extreme dynamics if we had provided participants with the number of listens rather than the number of downloads.

These speculations, and they are just speculation, are ultimately testable with additional experiments employing the "multiple-worlds" design used throughout this dissertation.

Chapter 7

Conclusions

Instead of reviewing what has already been covered in the previous chapters, we would like to conclude by looking forward and describing some possible next research projects. These projects are all in some way to the ideas presented in this dissertation and are all very preliminary. Many will never occur.

7.1 Musiclab variations

There are a number of very direct variations of musiclab that could be performed. For example, in chapter 6 we speculated about how the form of feedback provided to subjects would affect the inequality, unpredictability, and overall throughput of the market.

One common interpretation of the experiments results, which is slightly incorrect, is that when participants could see what other people "liked," which particular song became most "liked" was unpredictable. The problem with this interpretation is that participants did not see the ratings of other participants and our outcome measure was not based on unpredictability of ratings. Rather participants saw the previous download decisions of others and the unpredictability of downloads (or more specifically market share) was our measure of interest. To address the question of whether presenting people with ratings (rather than downloads) would produce the same effects we propose a "musiclab" of videos.[1] The design would be exactly the same except that since videos can't be downloaded the natural feedback to participants would be the average rating. Given the popularity of `www.youtube.com` and other video sites, and the common use of ratings as feedback on these sites, this experiment would seem natural to participants.[2]

An additional variation of musiclab which could be explored would be one that removed the listen decision. That is, participants would come to the site, but instead of choosing what to listen to, we would choose for them (either randomly or more popular songs more frequently). This would allow us to cleanly study social influence on the rating and download decision. However, removing the listening decision might seem unnatural to participants.

7.2 Micro-macro experiments and simulations

We performed four experiments, but these represent just a sample of the possible experiments that could have been performed within this framework. What if the there were 48,000 songs instead of 48? What if the songs were more (or less) homogeneous in appeal? And so on. In other words, how can we use the results from these four experiments to make statements about the outcomes that would be observed in the space of all possible experiments of this type?

One approach would involve statistical analysis of these experiments in order to estimate individual behavioral parameters which would then be used to calibrate simulations. If these simulation were able to successfully reproduce observed collective

[1] At the time musiclab began streaming video on a large-scale was not possible, but it is certainly possible today.

[2] Art might also work well in this format.

dynamics, they could then be used to explore the possible results of experiments that we did not perform. While this approach is promising, it is important to point out that it is not without problems. For example, if we find that participants are influenced by popularity at a given rate when there are 48 songs, how do we know that the same thing will happen when there are 48,000 songs? Maybe this behavioral parameter will change. In the end, the only way to know for sure if this simulation approach will be successful is to make predictions based on the simulations and then test them with new micro-macro experiments.

7.3 Where is "quality"?

In 1977, Dr. Sherman Lee, director of the Cleveland Museum, paid one million dollars for the painting *St. Catherine* by Matthias Grunewald. Later rumors began to spread that the painting was not actually painted by Grunewald, but was instead a forgery. Concerned by these rumors, Dr. Lee ordered pigmentation tests on the painting, tests which ultimately proved beyond a doubt that the painting was a fake (Alsop, 1986). What happened next is what's interesting.

First, the painting was removed from public display (Lee, 1986). Remember, this was a painting for which the museum was so excited that it paid one million dollars, and the physical painting had not changed in any way. Further, people actually saw the painting in a new way. Dr. Lee said that after discovering that it was a forgery the painting, "just fell to pieces," an apparently common reaction to paintings that are discover to be forgeries (Alsop, 1986).[3]

But is Dr. Lee telling the truth? Did how he saw the painting really change? Evidence to support his claim comes from a recent study involving Coke, Pepsi, and

[3]For another example of the role of authorship on artistic impression consider the case Rembrandt whose body of work is now in dispute, so much so that the Metropolitan Museum of Art in New York even had a show called "Rembrandt/Not Rembrandt" (von Sonnenburg, 1995; Liedtke et al., 1995).

and fMRI machine. When drinking Coke subjects were found to have more neural activity that is thought to indicate pleasure than when they drank an anonymous soft drink (which was either Coke or Pepsi). In other words, knowing they were drinking Coke actually made them like the same soda more. Thus, part of the pleasure of a Coke seems to be somewhere outside of the Coke itself, and part of the pleasure of looking at a Grunewald, likewise, seems to be outside of the painting itself.[4]

What are these things outside of the objects themselves and how do they come to be social constructed? We realize that this is not a well-framed questions, but we suspect that there is something interesting happening here. One attempt to begin to measure these things is with a series of pair-choice experiments (Bradley, 1976; Fienberg and Larntz, 1976; David, 1988; Dittrich et al., 1998; de Vries, 1998; Silverstein and Farrell, 2001). Participants would be presented with a question and two different images; for example, "Which is the better painting?" along with an image of two paintings. In some conditions, participants would choose without any additional information, but in other conditions participants would see the name of the artist, price at recent auction, or museum that owns the painting. To the extent that participants choices change with this additional information we can determine which type of information affects participants' choices most strongly, and further which pieces within those categories have the biggest effect. For example, participants may be more swayed by information about artist than by information about museum, but within the museum category participants maybe more affected by the Louvre than the M.O.M.A. Further, this information may affect different questions differently. For example, the information that might affect choices for "What is the better painting?" might not be the same as the information that affects choices for "Which painting would you want in your living room?" These primitive experiments might be a first step in helping to clarify the problem and develop some basic results from which to

[4]Similar results without brain scans have been found for peanut butter (Hoyer and Brown, 1990) and music (North et al., 2003; Colley et al., 2003).

attempt to construct a more general theory.

7.4 Parody

Cumulative advantage in cultural markets is clearly more complicated than just people being attracted to more popular things; it operates via numerous mechanism. Further research could explore the workings of these different mechanism and estimate there relative strengths in different markets.

The process of parody is a perfect example of cumulative advantage. An artist chooses to parody a work because it is already popular (why parody something that nobody knows?), and the parody, in turn, makes the original more popular. An excellent example of this process is the *Mona Lisa* (Sassoon, 2001, 2006). When creating *L.H.O.O.Q.*, the famous mustached parody, Marcel Duchamp choose to parody *Mona Lisa* because the painting was already popular, and this parody only reinforced the fame of the original painting, as well as set-off a cascade of other fame-reinforcing parodies (figure 7.1).[5] Biel (2005), in his case-study of *American Gothic* by Grant Wood, also emphasized the role of parody in the social construction of cultural icons. A better understanding of parody would aid in the understanding of cumulative advantage in cultural markets.

7.5 Friendcast

In cultural markets information about products spreads from person to person, but we have very little information about how this processes actually occurs (Watts and Peretti, 2007). For example, what kinds of content are most likely to spread?

[5]The full story of the rise of the Mona Lisa to global icon is an interesting one and in many ways consistent with the arguments made in dissertation. For an excellent history see Sassoon (2001, 2006).

(a) Mona Lisa (Da Vinci) (b) Mona Lisa (Duchamp) (c) Mona Lisa (Unknown)

Figure 7.1: Examples of parody of the Mona Lisa.

And by whom? Can this spreading be predicted or engineered?

Rather than relying on cumulative adoption curves, as in the Bass model (Bass, 1969; Bass et al., 1994), answering many of these questions require explicit individual-level data. We have begun a new research project with Sharad Goel to collect this information. Our website, `www.friendcast.org`, allows initiators to post content (text, audio, or video) which they then try to spread. The initiator is given a unique url which they can send to their friends. Each of these friends who visits the site is given a new unique url which they can pass to their friends, and so on. Because each url is unique, we can collect precise information about the chain of recruitments that begin from the initiator.[6] Figure 7.2 presents a screenshot of the experiment. A key feature of the site is that initiators, as well as subsequent chain members, can

[6]In some ways this tracking system is similar to the one used in respondent-driven sampling where participants pass uniquely numbered coupons (Heckathorn, 1997, 2002; Salganik and Heckathorn, 2004). The system also bears some resemblance the "chain tag" approach of Dodd (1956). Because it is electronic rather than paper-based, Friendcast can avoid many of the limitations of these paper-based approaches.

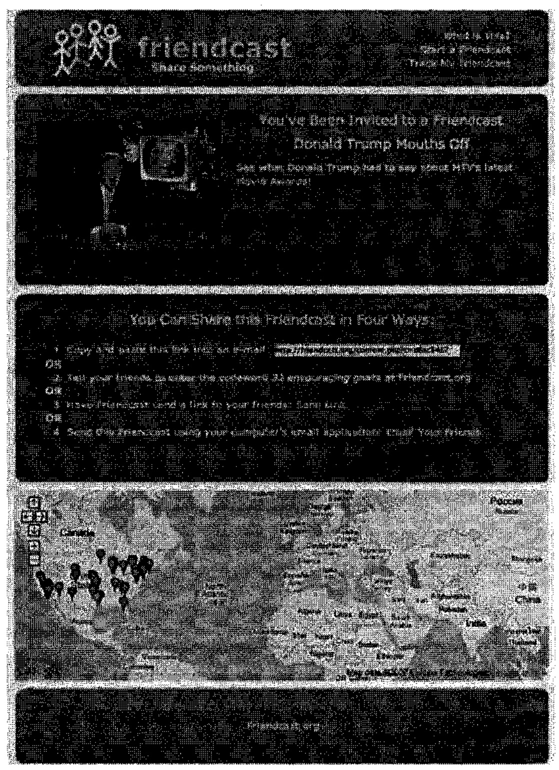


Figure 7.2: Screenshot from the Friendcast website.

monitor the spread of their content both geographically and in terms of waves. It is hoped that this ability to see how things spread, which is not available on other sites like `www.youtube.com`, will motive people to participate.

Some data of this form has been collected previous using the Forward Track framework[7] developed by Eyebeam (Watts and Peretti, 2007), but this data collection method did not allow regular people to submit content to be spread. We hope that having user-submitted content will allow for more extensive data-collection.

[7] `http://forwardtrack.eyebeamresearch.org/`

7.6 Museum tracking

One area where social influence may operate on choices is inside of art museum. We propose that when people enter an art museum, they would fill out a brief survey and then be given an electronic device which will track their physical path through the space.[8] Similar to the design of the crowd study of Milgram et al. (1969), a randomly sized "stimulus crowd" could be randomly placed in front of randomly chosen paintings. We could then measure how this crowd of confederates affects the behavior of other visitors. Do more people stop to examine the painting? Do they look at it for longer periods of time? Further, are some people or paintings more susceptible to these crowd effects? And how does these effect differ across museums?[9]

7.7 How do best-seller lists drive sales?

As was recently demonstrated in Sorensen (2007), incorrectly appearing on the New York Times best seller list improves book sales (see also Stern (1995)). However, there are many mechanisms through which appearing on the best-seller list could increase sales. One possible way to separate some of these is through a field experiment.

Many music and book stores have a special shelf which presents the current best-sellers, usually in order. During the course of this dissertation, we noticed that a local music store had a similar display, but rather than being based on national best-sellers, it was based on best-sellers at the particular store. Given that the list

[8]Although electronic measurement would not be required, this would significantly improve the resolution of the data over human observation (Klein, 1993). An alternative mechanical approach which could be done without the knowledge of visitors is the Hodometer (Bechtel, 1967). In this approach, researchers built weight sensors into the floor and so are able to track the aggregate pattern of visitors. Unfortunately, this approach does not allow researchers to explore the relationships between individual-level behavior and demographics.

[9]Even in the absence of experimental manipulation, the observational data could be used to determine class-based differences in art appreciation (Bourdieu, 1984).

was based on local popularity, consumers would have no way of knowing if the information was manipulated. We approached an associate who worked at the store about the possibility of randomly assigning different CDs to different rankings on different days.[10] Through random assignment we could isolate this point-of-sale popularity effect, independent of the overall popularity of the product. For example, how much would sales of the 10th most popular CD increase if people thought it was the most popular? How would the 5th most popular CD be affected by this same manipulation? What if these CDs were assigned to be the second most popular rather than the most popular? This experiment did not proceed, however, because sales at the store were very low making it difficult to detect any effect. In addition to this practical problem, this experiment may also raise ethical issues.[11]

7.8 Fashion game

One of the limitations of the musiclab experiment is that participants only made decisions at one time-point and could not change their decision based on the decisions of future participants. This leads to much simpler dynamics than in real social interaction where people are continually adjusting their behavior in reaction to the behavior of others who are also constantly adjusting their behavior and so on. One place where this repeated interaction framework is most important is in fashion. Every morning people must get dressed and decisions about what to wear are based on the decisions of others. How is it that "trucker hats" can become popular, but then fade away?

To explore the dynamics of fashion, we propose to develop a web-based iterated

[10]Rather than reshuffling the best-sellers, another approach would be to randomly place some objects onto the best-seller rack.

[11]A variant which does not involve manipulation, and therefore doesn't raise ethical concerns, would be to provide shoppers both national and local best-seller information and then somehow determine which has a greater effect on purchase decisions.

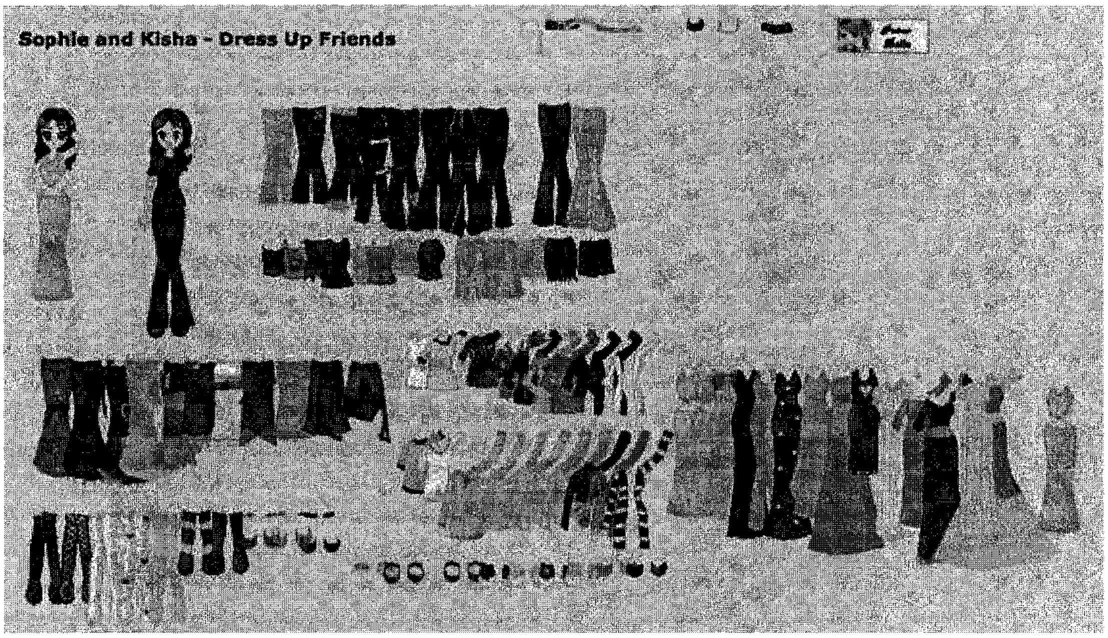


Figure 7.3: Screenshot from the Sophie and Kisha dress-up website.

"e-paper doll game." Many websites exist where people come and "dress up" a doll in different types of clothing (figure 7.3). We propose using this framework as a basis for a game. In this game several people would play simultaneously, each with their own doll. First, each participant would dress their doll without knowledge of how other participants will dress their dolls. After all dolls have been dressed, there would be a "fashion show" where each doll would be rated by the other participants. Participants would then re-dress their doll and have another fashion show. The experiment would proceed in this way for a given number of rounds.[12]

Several features of the interaction could be varied to explore the effects on fashion dynamics. For example, one could imagine three different information conditions: no rating information, individual rating information (i.e. participants would

[12]Logistically, this fashion game might be more difficult than musiclab because it would require participants to be online at the same time. However, the rounds would not need to happen in real-time and could instead take place over days, as is mail-based chess.

only see their own score in the fashion show), and global information (i.e. participants would see all the ratings). We suspect that fads would occur more dramatically in the higher information conditions. Another parameter to explore is the number of kinds of clothes that are available. We suspect that as the number of available choices decreases, the fads will become more pronounced. A final parameter to be explored is group size. We suspect that as the groups get larger, the fads will occur more dramatically. All of these experiments would take place in the multiple-worlds framework, and could therefore explore the arbitrariness of fashion.

7.9 Cumulative advantage and disadvantage in the life course

Cumulative advantage and disadvantage processes are at work outside of cultural markets as well. Consider the case of Dave who worked at the Firestone tire plant in Decatur, Illinois, for 20 years until his factory closed and he was laid-off (Sered and Fernandopulle, 2005). Losing a job is a terrible experience, but for Dave the loss was exacerbated because unemployment meant the loss of his health insurance, and with it, the means of paying for his heart medication. Without the medication, his health deteriorated and his relationship with his wife became strained. Because of the accumulation of these problems it was almost impossible for Dave to find a new job; he was stuck in what scholars Sered and Fernandopulle (2005) call the "death spiral." Dave's story illustrates the processes of cumulative disadvantage where problems beget more problems.[13]

But cumulative effects can also work in the opposite direction, creating new

[13]A similar account of spiral into homelessness in presented in Snow and Anderson (1993). On page 270 the authors write, "When we turn to [the life histories of the homeless], it becomes strikingly evident that although homelessness results from the confluence of various misfortunes, that confluence does not happen mysteriously. Rather, one misfortune or problem triggers or exacerbates another, which, in turn, leads to yet another until the individual is forced into the streets."

advantages from a single benefit. Between 1962 and 1967, 128 disadvantaged children in Ypsilanti, MI took part in the Perry Program to study the efficacy of early interventions in the lives of disadvantaged children. Half of the group was randomly assigned into a treatment condition, where they received two and half hours of preschool each weekday morning and weekly home visits from a teacher. This early intervention led to gains for the children which started a process of cumulative advantage, where success lead to future success, and these gains continued long after their preschool was finished. Follow-up data collected 40 years later showed that children in the treatment group scored higher on achievements tests, earned higher wages, were more likely to own a home, and were less likely to go on welfare or become incarcerated (Schweinhart et al., 2005; Heckman, 2006).

Both these stories, each in their own way, demonstrate the way that exogenous events (either positive or negative) can push people onto new life trajectories, which then become self-sustaining. Unfortunately, social scientists do not really understand the mechanisms involved in these virtuous and vicious cycles (DiPrete and Eirich, 2006). With a fuller understand of cumulative advantage and cumulative disadvantage, policy makers could design programs to make it less likely that people such as Dave get stuck in a death spiral. Moreover, policy makers could design better interventions like the Perry Program which set off a process of cumulative advantage.

We suspect that progress in this area may be made by modeling within the dynamics systems perspective which allows for models of coupled differential equations. This approach, unlike those used traditionally by social scientists, is explicitly dynamic, a characteristic need for modeling cumulative processes, and allows for non-linear effects, where a small change can make a big difference and a big change can make a small difference. Further, this approach naturally allows for the modeling of the interdependency among different life contexts (e.g. work, health, family), an under-studied and potential important feature of cumulative effects in the life course (DiPrete and Eirich, 2006). A better understanding the dynamic interdependency among life contexts will provide a fuller picture of the processes which generate social inequality.

These stylized models will yield results that will be just that, stylized. Therefore, we hope the modeling will yield new insight which can be combined with the analysis of existing longitudinal data, including the Panel Study of Income Dynamics (PSID) which collected longitudinal data on the health, economic, and social behaviors of about 8,000 families from 1968 to present. Another possibly interesting source of data is the Swedish Twin Registry (STR), which has data on monozygotic (identical) twins who are raised in the same household. Ordinarily, twin research involves studying monozygotic twins raised apart or adopted siblings raised together, but by studying the trajectories of monozygotic twins raised together, we can, as much as possible, "control" for background and study the effects that specific exogenous life events have on future outcomes. Current studies which rely on statistical controls to account for individual heterogeneity are often unconvincing (for example Elman and O'Rand (2004)).

We hope that this new approach to life-course research will both further the social science literature and suggest new policy approaches. For example, many downward spirals in the United States could be prevented by decoupling health insurance and work, thus preventing problems in one context from spreading to the other. This idea of decoupling might suggest complete decoupling of all contexts, but that would also limit the possibility for some contexts to provide buffers for shocks that occur in other contexts. Decoupling would also remove the possibility of upward spirals. Ideally we would like to design a system with coupling that allows for virtuous cycles, but not vicious ones. Also, research in this vein may help us understand how some interventions in one context can have knock-on effects in other contexts. For example, programs designed to promote health in children may also improve educational performance. These indirect effects may be easier to generate than direct effects targeted just at educational performance. Ultimately, the design of efficient and effective social programs will be aided by a better understanding of cumulative advantage in the life course.

7.10 Final thoughts

As can be seen by the proceeding discussion, this dissertation can lead in many different research directions. We suspect, however, that the evolution of this particular research trajectory will be unpredictable.

Appendix A

Appendix to chapter 2

Experiments 1, 2, and 3 were approved by the Columbia University Institutional Review Board as Protocol IRB-AAAA5286. Experiment 4 was approved as Protocol IRB-AAAB1483. We used the following forms for subject consent (A.1.1), assent (A.1.2), and parental/guardian consent(A.1.3). Also, as approved by the IRB, we used the following email to the bands (A.2) and band consent form (A.3). The text that was used in the experiment is presented in section A.4 and the survey that was given to all participants is presented in section A.5.

A.1 Consent forms

A.1.1 Adult consent form

```
Investigators:
--------------
Prof. Duncan J. Watts (1)
Matthew J. Salganik (1)
Dr. Peter S. Dodds (2)
(1) Department of Sociology, Columbia University
```

(2) Institute for Social and Economic Research and Policy
Contact: <consent@cdg.columbia.edu>

Investigators' Statement:

We are asking you to participate in a research study funded by the Institute for Social and Economic Research and Policy (ISERP) at Columbia University. The purpose of this consent form is to give you the information you will need to help you decide whether or not to take part in the study. If, after reading the information provided, you have further questions regarding the experiment please contact us at <consent@cdg.columbia.edu>

The purpose of this study is to understand how people evaluate music. You will be presented with a list of sound recordings ("songs"). You can listen to a song by clicking on the appropriate link. Each time you listen to a song, we will ask you a few questions about your opinion of the song. Finally, after doing this evaluation, you will have a chance to download the song.

All songs used in this experiment are under copyright protection. We have obtained permission from the appropriate copyright holder to use them in this research. It is your responsibility to observe the copyright restrictions on this music. Once downloaded, you can listen to a song anytime you like. For your personal use only you may: save it to your hard drive, burn it to a CD, or transfer it to your portable mp3 player.

The personal information you provide will be securely protected. Once you have agreed to participate in this research project, we will ask you for some information about yourself in order to help us understand how demographic factors like age, ethnicity, and musical taste can be used to understand the behavior of people in the project.

The total amount of time needed for this experiment ranges from about 5 minutes to more than an hour depending on the number of songs to which you choose to listen.

You may terminate your involvement in the experiment at any time by visiting the <XXX> section of the site and following the Delete User Profile link. We may contact you again to ask if you would like to see the final results of the experiment or participate in future experiments.

Institutional Review Board Statement:

Participation in this research project is entirely voluntary. You may refuse to participate or withdraw from participation at any time without jeopardy to future medical care, employment, student status or other entitlements. The investigator may withdraw you at his/her professional discretion.

If, during the course of the study, significant new information that has been developed becomes available which may relate to your willingness to continue to participate, the investigator will provide this information to you. Any information derived from this research project that personally identifies you will not be voluntarily released or disclosed without your separate consent, except as specifically required by law.

If at any time you have questions regarding the research or your participation, please email to: <consent@cdg.columbia.edu>

If at any time you have comments regarding the conduct of this research or questions about your rights as a research subject, you should contact the Institutional Review Board Administrator at +1-212-870-3585 or <askirb@columbia.edu>.

Statement of Benefits and Costs to the Subject:

We can foresee no risks to the subjects from this research. Subjects will benefit from a chance to hear and download new music.

Subject's Statement:

This study has been explained to me. I am at least 18 years old and I volunteer to take part in this research. I have had the opportunity to ask questions. If I have questions later on about the research I can ask the investigators listed above. If I have questions about my rights as a research subject, I can contact the Institutional Review Board at +1-212-870-3585 or <askirb@columbia.edu>.

A.1.2 Assent form

Investigators:

Prof. Duncan J. Watts (1)

Matthew J. Salganik (1)

Dr. Peter S. Dodds (2)

(1) Department of Sociology, Columbia University

(2) Institute for Social and Economic Research and Policy

Contact: <consent@cdg.columbia.edu>

Investigators' Statement:

We are asking you to participate in a research study funded by the Institute for Social and Economic Research and Policy (ISERP) at Columbia University. The purpose of this assent form is to give you the information you will need to help you decide whether or not to take part in the study. If, after reading the information provided, you have further questions regarding the experiment please contact us at <consent@cdg.columbia.edu>

The purpose of this study is to understand how people evaluate music. You will be presented with a list of sound recordings ("songs"). You can listen to a song by clicking on the appropriate link. Each time you listen to a song, we will ask you a few questions about your opinion of the song. Finally, after doing this evaluation, you will have a chance to download the song.

All songs used in this experiment are under copyright protection. We have obtained permission from the appropriate copyright holder to use them in this research. It is your responsibility to observe the copyright restrictions on this music. Once downloaded, you can listen to a song anytime you like. For your personal use only you may: save it to your hard drive, burn it to a CD, or transfer it to your portable mp3 player.

The personal information you provide will be securely protected. Once you have agreed to participate in this research project, we will ask you for some information about yourself in order to help us understand how demographic factors like age, ethnicity, and musical taste can be used to understand the behavior of people in the project.

The total amount of time needed for this experiment ranges from about 5 minutes to more than an hour depending on the number of songs to which you choose to listen.

You may terminate your involvement in the experiment at any time by visiting the <XXX> section of the site and following the Delete User Profile link. We may contact you again to ask if you would like to see the final results of the experiment or participate in future experiments.

Institutional Review Board Statement:

Participation in this research project is entirely voluntary. You may refuse to participate or withdraw from participation at any time without jeopardy to future medical care, employment, student status or other entitlements. The investigator may withdraw you at his/her professional discretion.

If, during the course of the study, significant new information that has been developed becomes available which may relate to your willingness to continue to participate, the investigator will provide this information to you. Any information derived from this research project that personally identifies you will not be voluntarily released or disclosed without your separate assent, except as specifically required by law.

If at any time you have questions regarding the research or your participation, please email to: <consent@cdg.columbia.edu>

If at any time you have comments regarding the conduct of this research or questions about your rights as a research subject, you should contact the Institutional Review Board Administrator at +1-212-870-3585 or <askirb@columbia.edu>.

Statement of Benefits and Costs to the Subject:

We can foresee no risks to the subjects from this research. Subjects will benefit from a chance to hear and download new music.

Subject's Statement:

This study has been explained to me. I am under 18 years old and I volunteer to take part in this research. I have had the opportunity to ask questions. If I have questions later on about the research I can ask the investigators listed above. If I have questions about my rights as a research subject, I can contact the Institutional Review Board at +1-212-870-3585 or <askirb@columbia.edu>.

A.1.3 Parent/guardian consent form

Investigators:

Prof. Duncan J. Watts (1)

Matthew J. Salganik (1)

Dr. Peter S. Dodds (2)

(1) Department of Sociology, Columbia University

(2) Institute for Social and Economic Research and Policy

Contact: <consent@cdg.columbia.edu>

Investigators' Statement:

We are asking your child to participate in a research study funded by the Institute for Social and Economic Research and Policy (ISERP) at Columbia University. The purpose of this consent form is to give you the information you will need to help you decide whether or not your child should take part in the study. If, after reading the information provided, you have further questions regarding the experiment please contact us at <consent@cdg.columbia.edu>

The purpose of this study is to understand how people evaluate music. Your child will be presented with a list of sound recordings ("songs"). Your child can listen to a song by clicking on the appropriate link. Each time your child listens to a song, we will ask your child a few questions about her/his opinion of the song. Finally, after doing this evaluation, your child will have a chance to download the song.

All songs used in this experiment are under copyright protection. We have obtained permission from the appropriate copyright holder to use them in this research. It is your and your child's responsibility to observe the copyright restrictions on this music. Once downloaded, your child can listen to a song anytime she/he likes. For her/his personal use only she/he may: save it to her/his hard drive, burn it to a CD, or transfer it to her/his portable mp3 player.

The personal information your child provides will be securely protected. If your child participates in this research project, we will ask her/him for some personal information in order to help us understand how demographic factors like age, ethnicity, and musical taste can be used to understand the behavior of people in the project.

The total amount of time needed for this experiment ranges from about 5 minutes to more than an hour depending on the number of songs to which your child chooses to listen.

You may terminate your child's involvement in the experiment at any time by visiting the <XXX> section of the site and following the Delete User Profile link. We may contact your child again to ask if your child would like to see the final results of the experiment or participate in future experiments.

Institutional Review Board Statement:

Participation in this research project is entirely voluntary. You or your child may refuse to participate or withdraw from participation at any time without jeopardy to future medical care, employment, student status or other entitlements. The investigator may withdraw your child at his/her professional discretion.

If, during the course of the study, significant new information that has been developed becomes available which may relate to your child's willingness to continue to participate, the investigator will provide this information to your child. Any information derived from this research project that personally identifies your child will not be voluntarily released or disclosed without your separate consent, except as specifically required by law.

If at any time you have questions regarding the research or your child's participation, please email to: <consent@cdg.columbia.edu>

If at any time you have comments regarding the conduct of this research or questions about your child's rights as a research subject, you should contact the Institutional Review Board Administrator at +1-212-870-3585 or <askirb@columbia.edu>..

Statement of Benefits and Costs to the Subject:

We can foresee no risks to your child from this research. Your child will benefit from a chance to hear and download new music.

Subject's Statement:

This study has been explained to me. I am the parent/guardian of the child in question and I give my consent for my child to take part in this research. I have had the opportunity to ask questions. If I have questions later on about the research I can ask the investigators listed above. If I have questions about my rights as a research subject, I can contact the Institutional Review Board at +1-212-870-3585 or <askirb@columbia.edu>.

A.2 Email to bands

Subject Line: Question About Your Music

Hi,

My colleagues and I are doing some research about why certain music becomes popular. We heard about your band from the purevolume.com website and would like to use one of your songs in an experiment we are planning.

The experiment involves creating a webpage where people can listen to, rate, and download new music. It offers you a great chance to get some publicity for your work. We expect between 10,000 and 100,000 music fans from all over the world to visit our site and listen to the songs posted. If we use one of your songs, we will also include a link to your website so that interested fans can learn more about you.

Since your work is copyright protected, we need to get your permission before using it. Also, our university requires that we get your permission in writing. The university approved form is attached.

If you want to participate in the research, please:

1) respond to this email and attach the Consent Form;
2) provide the information requested at the end of the Consent Form, including the name of the song;
3) provide the address of your website where fans can learn more about you; and
4) attach a copy of one of your songs (or send us a URL where we can access it).

You can learn more about us, and our past work, by visiting our website http://cdg.columbia.edu. Thanks for your consideration and let me know if you have any questions. If you do not wish to receive emails from us in the future, please click here <URL here>.

Matt

--

Matthew Salganik
Department of Sociology
Columbia University
1180 Amsterdam Avenue
New York, NY 10027
Telephone: +1-212-854-0367
Email: mjs2105@columbia.edu
Website: http://www.columbia.edu/~mjs2105

A.3 Band consent form

Consent Form for Copyright Holders

Investigators:

Prof. Duncan J. Watts, Department of Sociology, Columbia University
Matthew J. Salganik, Department of Sociology, Columbia University
Dr. Peter S. Dodds, Institute for Social and Economic Research and Policy
Contact Information: bands@cdg.columbia.edu or (212) 854-0367

Investigators' Statement:

We are asking to use your sound recording ("song") for a research study funded by the Institute for Social and Economic Research and Policy (ISERP) at Columbia University. The purpose of this form is to give you the information you will need to help you decide whether or not to provide your song as part of the study. If, after reading the information provided, you have further questions regarding the experiment please contact us at bands@cdg.columbia.edu or (212) 854-0367.

The purpose of this study is to understand how people evaluate music and how that evaluation can be affected by peers. Participants in the experiment will be divided into two groups. Everyone will be offered a chance to listen to, rate, and download a set of songs, including yours. When answering questions about the songs, participants in one group will receive information about the behavior of others in the group. Participants in the second group will not get this information. By comparing the behavior of these two groups we hope to learn more about why songs become popular.

If you allow us to use your song, you will still maintain all copyright ownership of your work. Our research is academic and not commercial. We will not directly profit financially from this work in any way.

Human Subjects Review Committee Statement:

Participation in this research project is entirely voluntary. You may refuse to participate or withdraw from participation at any time without jeopardy to future employment, student status, medical care, or other entitlements. The investigator may withdraw your song at his professional discretion at any time and for any reason.

If, during the course of the study, new information becomes available, which may relate to your willingness to continue to participate, the investigator will provide this information to you.

If at any time you have questions regarding the research or your participation, please contact us at bands@cdg.columbia.edu or (212) 854-0367.

If at any time you have comments regarding the conduct of this research or questions about your rights as a research subject, you should contact the Human Subjects Review Committee Administrator at <askirb@columbia.edu> or (212) 854-1324.

Statement of Benefits and Risks:

This experiment will offer you free publicity for your song. We expect between 10,000 and 100,000 music fans from around the world to participate in the experiment. These people may listen to and download your song. We will also provide a link to the band's website so that interested participants can learn more about the music they enjoy.

There are potential risks to you from allowing one of your songs to be in our experiment. These risks are substantially similar to the risks involved in making the song available to the public in any form. Once the work is released, in any form, people may make unauthorized copies of the song. Once the material has left the webpage, we, as experimenters, have no control over what will happen to your song. However, we will remind study participants that the work they are listening to and downloading is copyright protected.

Participant's Statement:

This study has been explained to me. I am the copyright holder of the sound recording in question. I give you permission to post my song on your research website and use for future related research.

I have the right to grant this permission. I am 18 years of age or older.

I have had the opportunity to ask questions. If I have questions later on about the research, I may contact the investigators listed above. If I have questions about my rights as a research participant, I can contact the Human Subjects Review Committee at <askirb@columbia.edu> or (212) 854-1324.

Copyright Holder:

Date:

Signature [type the signer's name]:

Name:

Email:

Telephone:

Name of Song:

Investigator:

Date: May 15, 2004

Signature: [duncan j. watts]

Name: Professor Duncan J. Watts

Email: bands@cdg.columbia.edu

Telephone: (212) 854-0367

A.4 Website text

Below is the text used on the specific webpages at our site. Screenshots from the site are in figures 2.2 to 2.11.

A.4.1 Welcome screen

Music Lab is a research project conducted by scientists from Columbia University to learn about how people form opinions about music. If you participate in Music Lab you will have a chance to download free new music.

After answering a few questions about yourself, you will be presented with a menu of songs by cool new artists. Your participation will take between 5 minutes and about two hours depending on how many songs you choose to listen to.

If you understand and are ready to participate, please click on the appropriate button below. If you would like to learn more about the research, please investigate the links at the top of the page. Enjoy.

A.4.2 Instructions

The instruction screen had the following text which was also accompanied by a figure (A.1(a) or A.1(b) depending on the experimental condition).

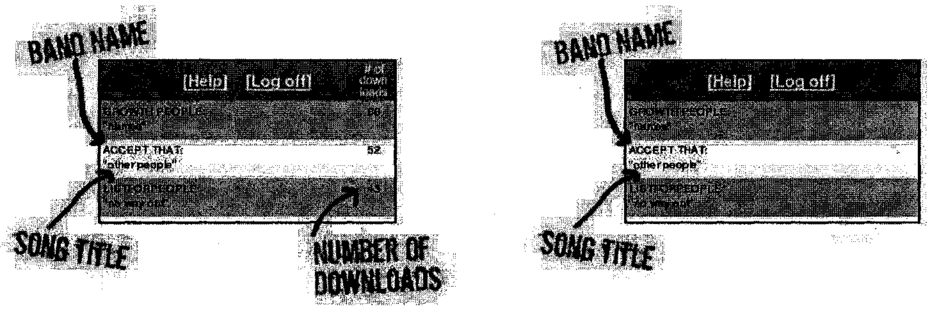
Thanks for completeing the survey. Now it's time to get to the music.

On the next screen you will be presented with a menu of songs. <In social influence condition: The number to right of each song is the number of times it has been downloaded previously.> Just click on the cell of the song you want to hear and it will start to play.

While it's playing you can rate it -- you don't need to wait until the end of the song. After your rating is submitted you will have a chance to download the song.

A.4.3 Log off

After participants clicked log off they were taken to a screen thanking them for participating and then providing them links to the webpages of the bands sorted into three categories: downloaded, listened only, and all others. The text of the screen was as follows:



(a) Social influence condition

(b) Independent condition

Figure A.1: Instructions image in social influence and independent condition.

Thanks for participating in the Music Lab. If you liked it, please tell your friends.

Below are links to information about the bands whose music you just heard. For more information just click on the song that interests you. Thanks again.

Downloaded song(s):

Listened only song(s):

The rest:

A.5 Survey

All participants had to complete a survey. Here are the questions that we used.

Background information:

In which country do you currently live?

<select from drop-down list>

If you live in the United States, please enter your zip code
<fill in the blank>

In what year were you born?
<fill in blank>

What is your gender?
<select from drop-down list:
male, female>

Compared to your circle of friends, how likely are you to be asked for advice about music?
<select from drop-down list:
much more likely, more likely, less likely, much less likely>

How did you hear about this experiment?
check all that apply
<web site or blog>
<internet ad>
<email from music lab>
<one of the bands>
<friend told me about a specific song>
<friend told me about a specific band>
<friend told me about the site in general>
<search engine>

<other>

Computer information:

What type of internet connection are you currently using?

<select from drop-down list:
broadband, dial-up>

Where are you while you are participating in this experiment?

<select from drop-down list:
office/work, home, public computer, other>

How would you rate your ability to use World Wide Web?

<select from drop-down list:
excellent, good, fair, poor, very poor>

In the past 30 days,how often have you visited a site for information about music or concerts?

<select from drop-down list:
never, 1-2 times, 3-5 times, more than 5 times>

Have you ever purchased a record as a result of hearing it on the web?

<select from drop-down list:
yes, no>

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

<fill in the blank>

Music Information:

How familiar are you with the following bands *:

Guys on Couch

<radio buttons:

don't know it at all, heard of it, know it pretty well>

Grover Dill

<radio buttons:

don't know it at all, heard of it, know it pretty well>

Peter on Fire

<radio buttons:

don't know it at all, heard of it, know it pretty well>

U2

<radio buttons:

don't know it at all, heard of it, know it pretty well>

Remnant Soldier

<radio buttons:

don't know it at all, heard of it, know it pretty well>

Please provide your email address so that we can tell you about the results of Music Lab:

<fill in the blank>

Appendix B

Appendix to chapter 3

B.1 Other presentations of success outcomes

Tables B.1 and B.2 present the top 10 songs in each world from experiments 1 and 2. One limitation of these tables is that they provide no information about the market share of the songs. Therefore, figures B.1 and B.2 present multiple-thermometer plots of all songs in all worlds, and figure B.3 present these previous figures side-by-side and on the same scale. Figures B.4 and B.5 present multiple-thermometer plots of the mean ratings of all songs in all worlds, and figure B.6 present these figures side-by-side for easy comparison. Finally, figures B.7 and B.8 present multiple-thermometer plots of the probability of download given listen (i.e., batting average) for all songs in all worlds, and figure B.9 present these figures side-by-side for easy comparison.

Experiment 1

Rank	Social Influence								Independent
	World 1	World 2	World 3	World 4	World 5	World 6	World 7	World 8	
1	parker theory	parker theory	parker theory	parker theory	parker theory	silent film	not for scholars	parker theory	parker theory
2	forthfading	star climber	selsius	shipwreck union	forthfading	parker theory	unknown citizens	by november	stunt monkey
3	stunt monkey	by november	hydraulic sandwich	ember sky	simply waiting	stunt monkey	stunt monkey	the fastlane	the fastlane
4	star climber	the fastlane	forthfading	the fastlane	not for scholars	benefit of a doubt	parker theory	shipwreck union	star climber
5	hartsfield	silent film	stunt monkey	unknown citizens	star climber	star climber	ember sky	unknown citizens	unknown citizens
6	the fastlane	hall of fame	miss october	silent film	beerbong	unknown citizens	by november	go mordecai	miss october
7	by november	shipwreck union	benefit of a doubt	silverfox	by november	the fastlane	miss october	ember sky	silent film
8	go mordecai	not for scholars	by november	hydraulic sandwich	stunt monkey	miss october	beerbong	forthfading	forthfading
9	not for scholars	hydraulic sandwich	not for scholars	stunt monkey	the fastlane	not for scholars	the fastlane	silent film	by november
10	benefit of a doubt	forthfading	silent film	selsius	silent film	ember sky	selsius	silverfox	hartsfield

Table B.1: Top 10 songs in each world in experiment 1. In the eight social influence worlds there were three different first place songs. It is also the case that Parker Theory was ranked first in six of the eight social influence worlds.

Experiment 2

Rank	Social Influence								
	World 1	World 2	World 3	World 4	World 5	World 6	World 7	World 8	Independent
1	parker theory	parker theory	52metro	parker theory	parker theory	the fastlane	silent film	parker theory	parker theory
2	stunt monkey	forthfading	forthfading	stunt monkey	the fastlane	parker theory	parker theory	forthfading	the fastlane
3	undo	silent film	parker theory	the fastlane	selsius	unknown citizens	forthfading	the fastlane	stunt monkey
4	unknown citizens	the fastlane	unknown citizens	52metro	stunt monkey	silent film	simply waiting	selsius	hall of fame
5	the fastlane	unknown citizens	shipwreck union	by november	by november	ryan essmaker	hydraulic sandwich	benefit of a doubt	forthfading
6	miss october	stunt monkey	the fastlane	miss october	forthfading	far from known	far from known	post break tragedy	miss october
7	silent film	star climber	undo	ember sky	hydraulic sandwich	52metro	hartsfield	stunt monkey	silent film
8	forthfading	hall of fame	selsius	silent film	silent film	simply waiting	unknown citizens	deep enough to die	go mordecai
9	shipwreck union	hartsfield	star climber	beerbong	undo	hall of fame	not for scholars	summerswasted	hartsfield
10	benefit of a doubt	the broken promise	sibrian	benefit of a doubt	benefit of a doubt	by november	selsius	hydraulic sandwich	unknown citizens

Table B.2: Top 10 songs in each world in experiment 2. In the eight social influence worlds there were four different first place songs. It is also the case that Parker Theory was ranked first in five of the eight social influence worlds.

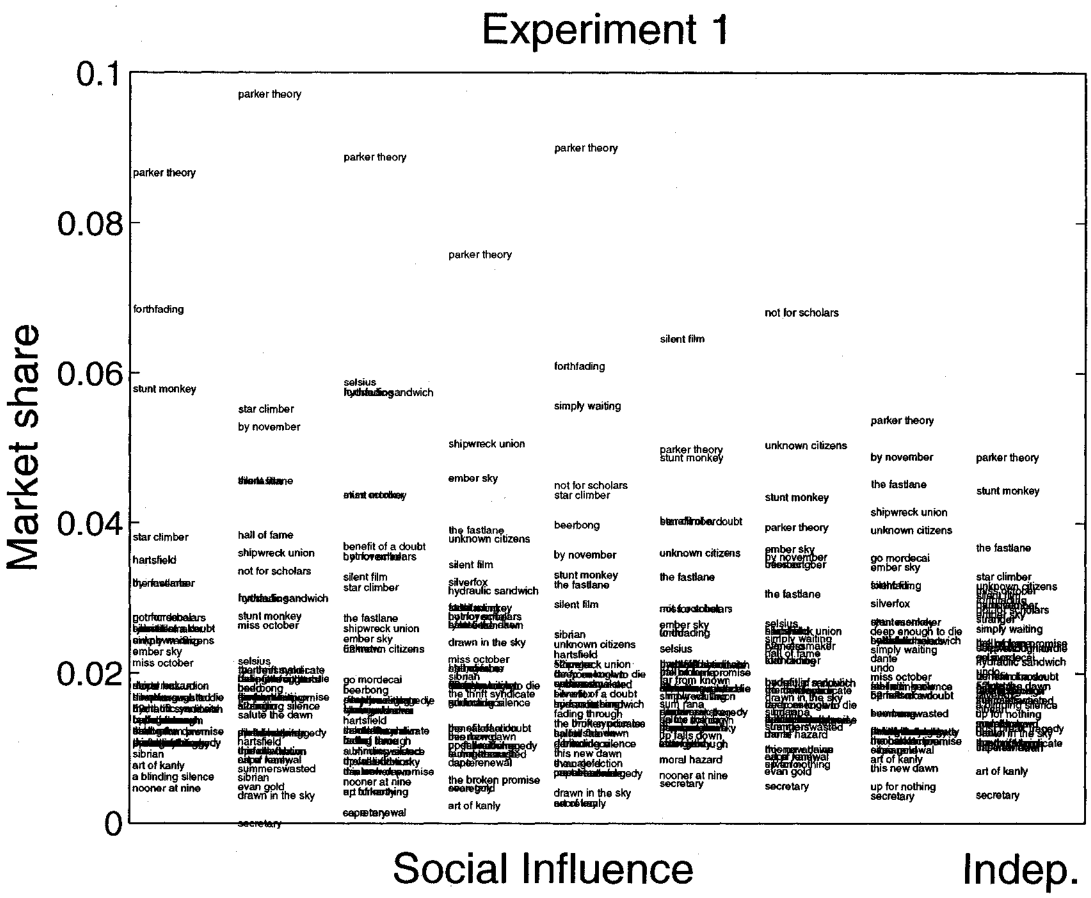


Figure B.1: Multi-thermometer plots of market share in experiment 1.

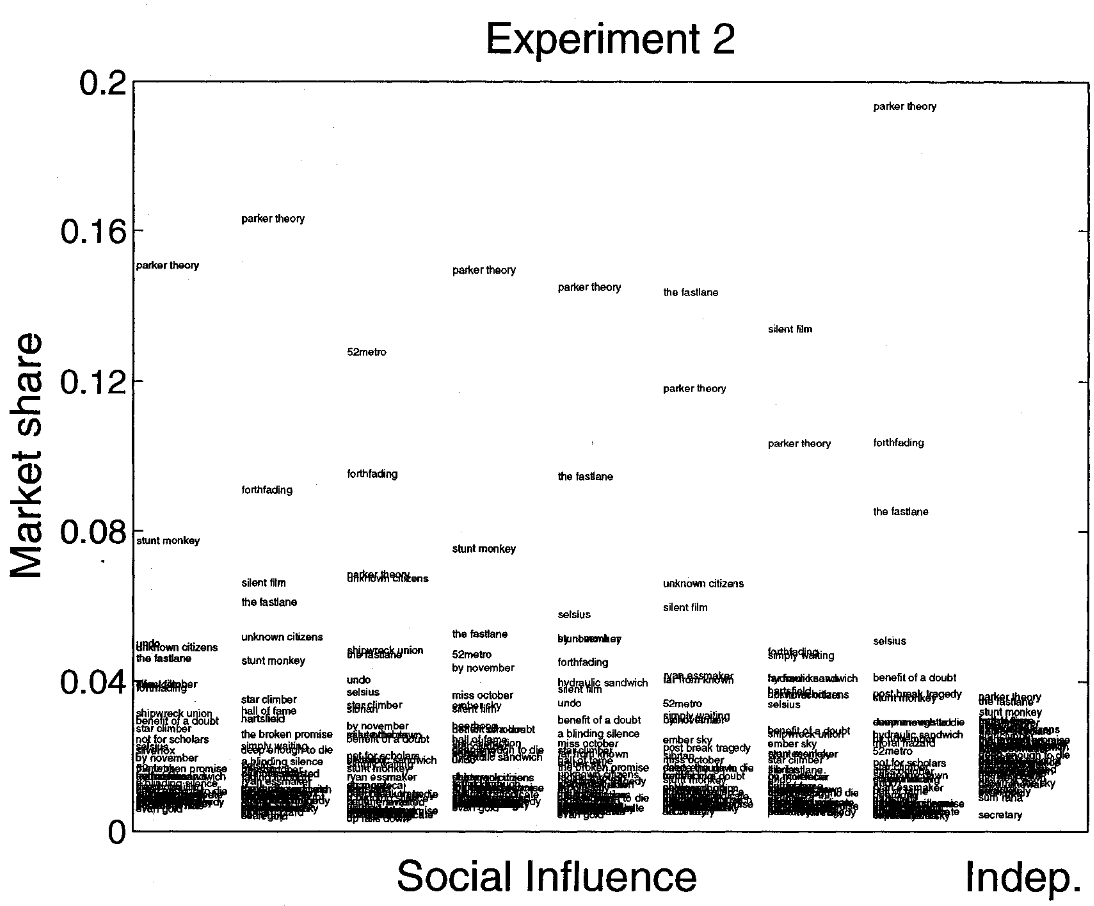
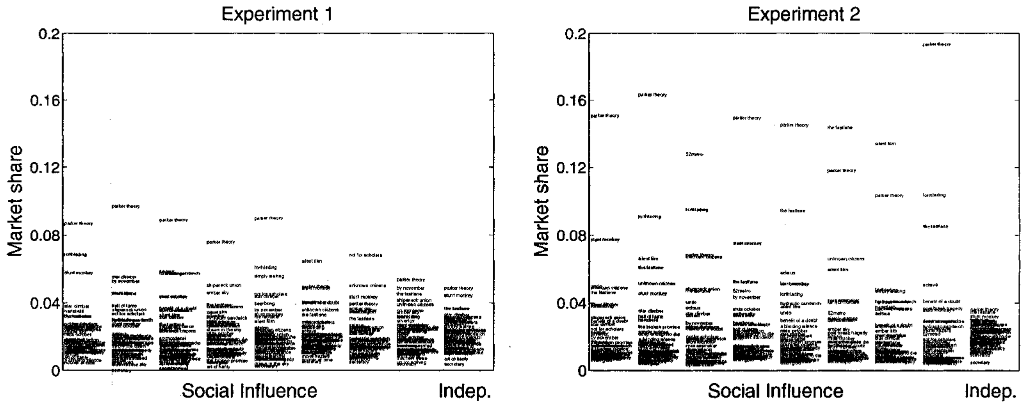


Figure B.2: Multi-thermometer plots of market share in experiment 2.



(a) Experiment 1

(b) Experiment 2

Figure B.3: Multi-thermometer plots of market share in experiments 1 and 2.

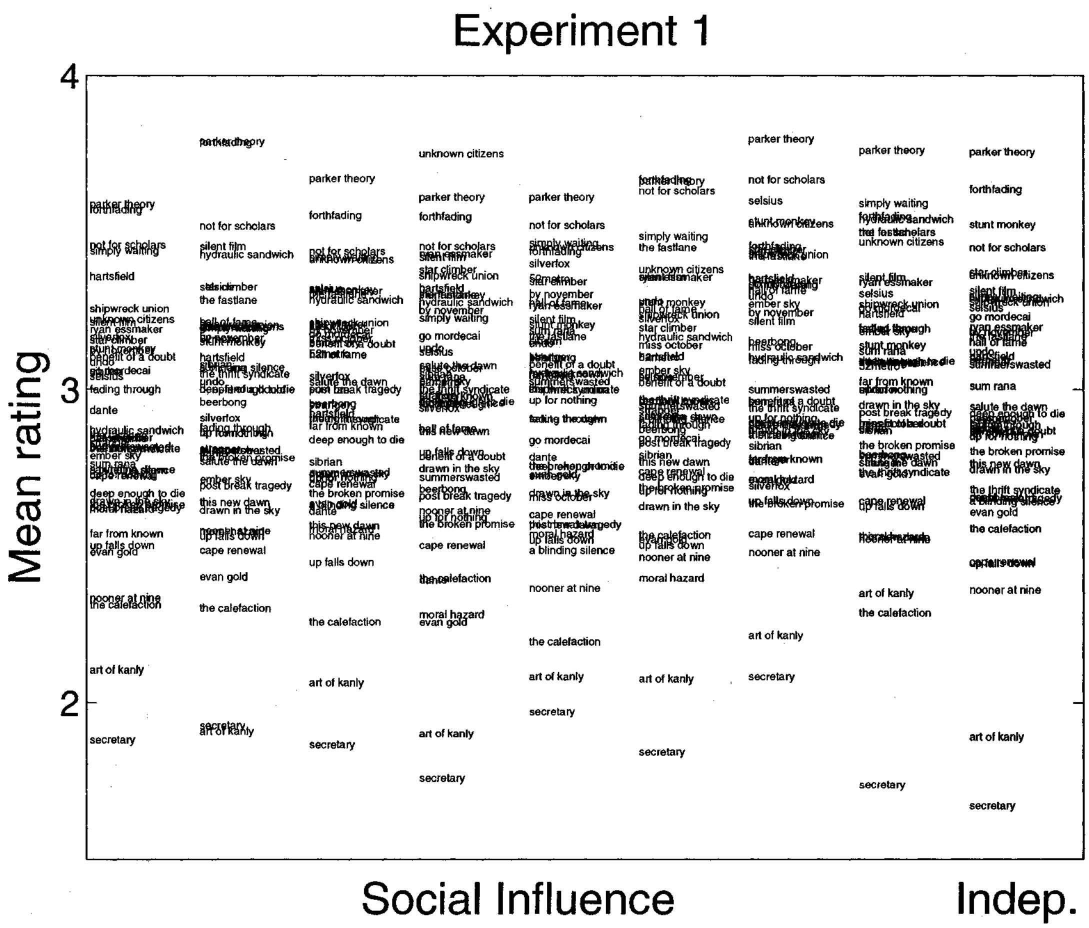


Figure B.4: Multi-thermometer plot of mean rating for each song in each world in experiment 1.

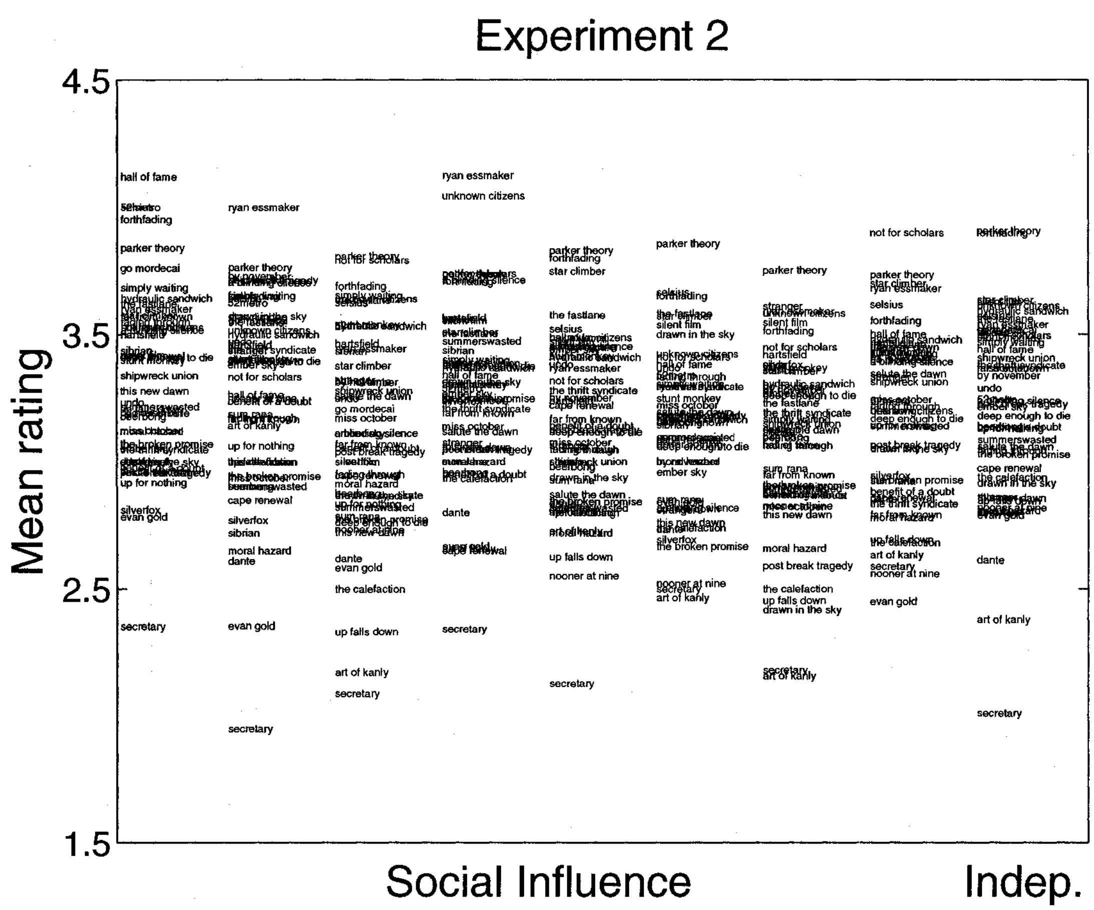
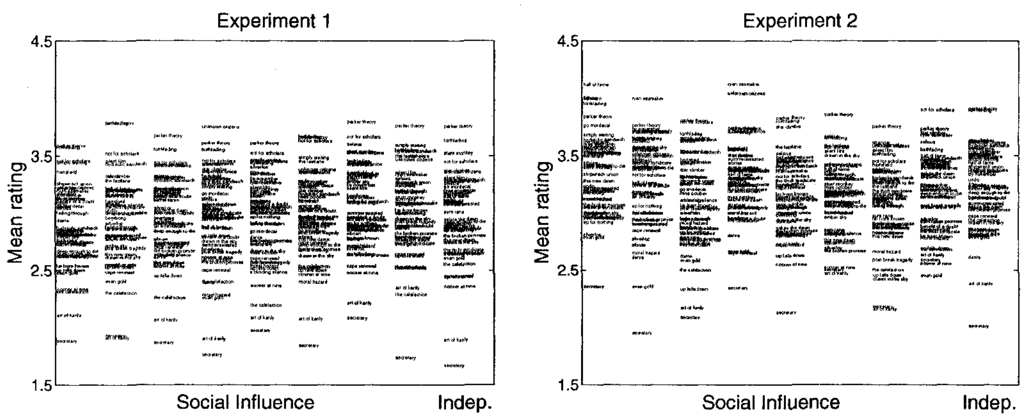


Figure B.5: Multi-thermometer plot of mean rating for each song in each world in experiment 2.



(a) Experiment 1

(b) Experiment 2

Figure B.6: Multi-thermometer plot of mean rating for each song in each world in experiments 1 and 2.

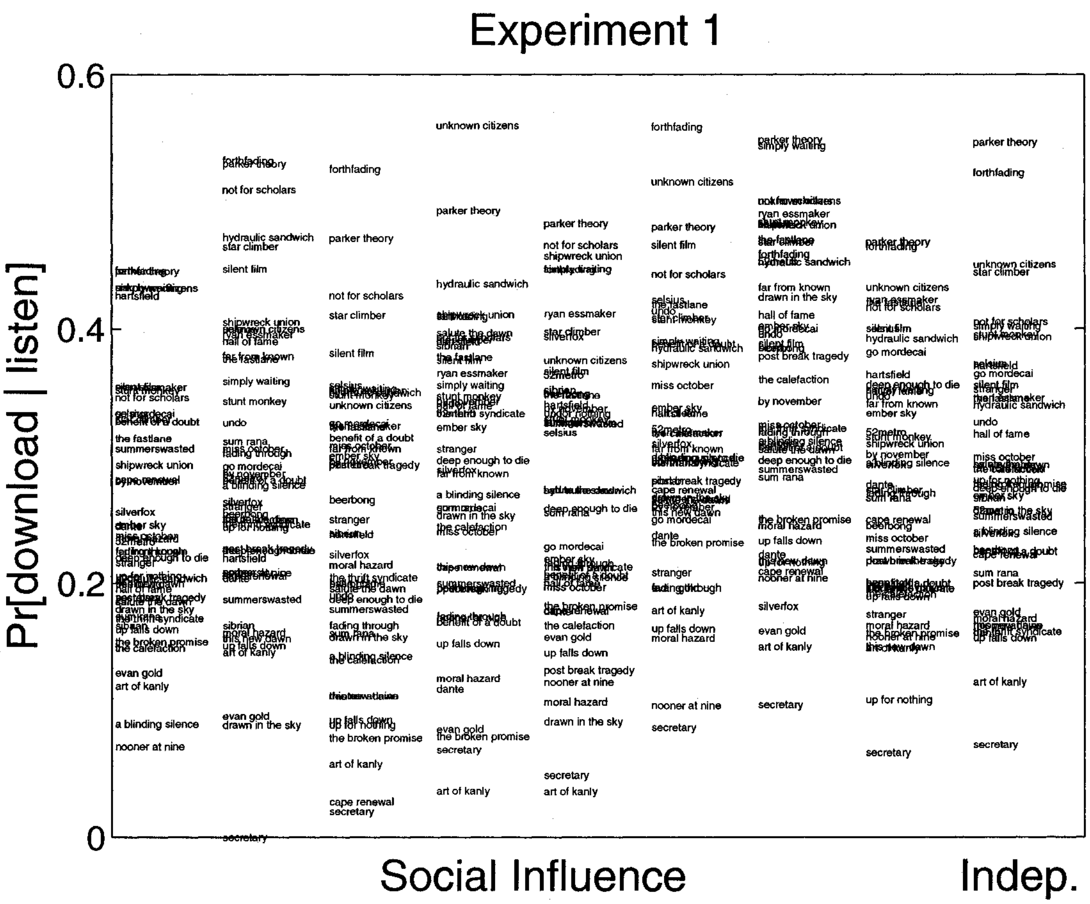


Figure B.7: Multi-thermometer plot of the probability of download given listen for each song in each world in experiment 1.

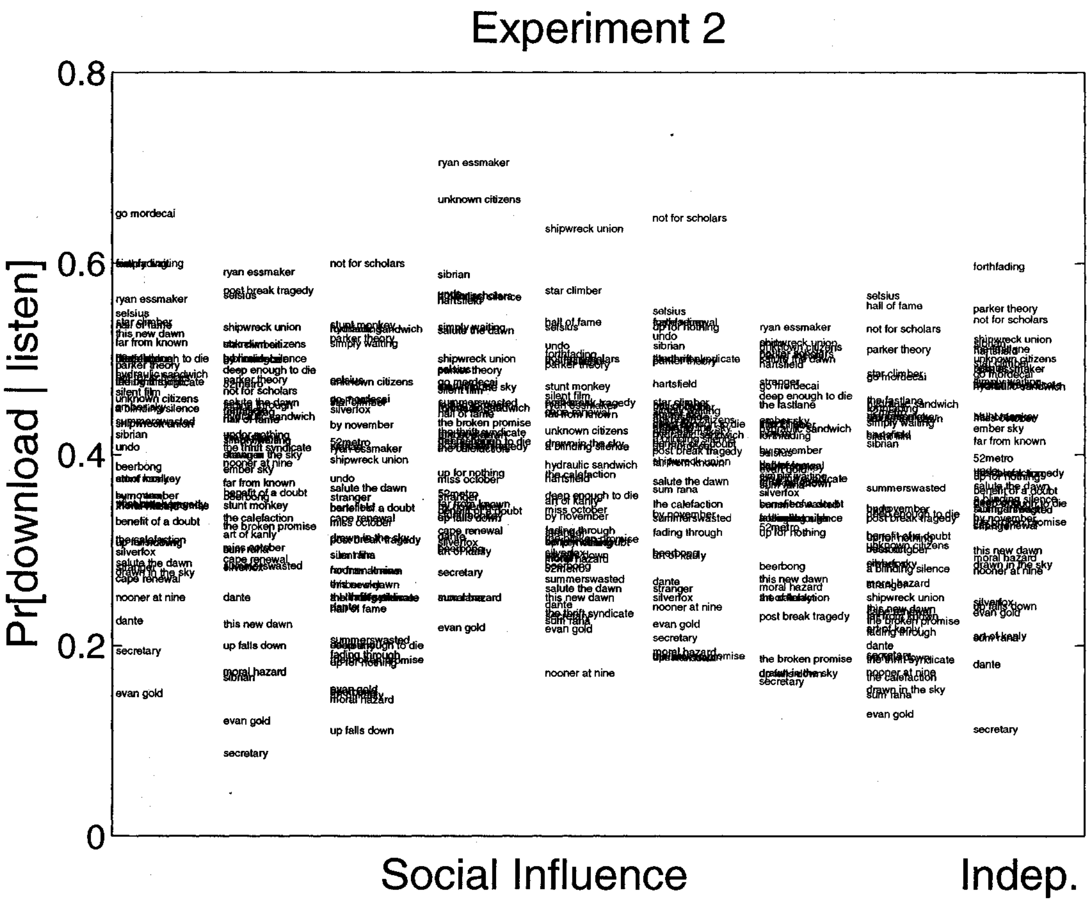
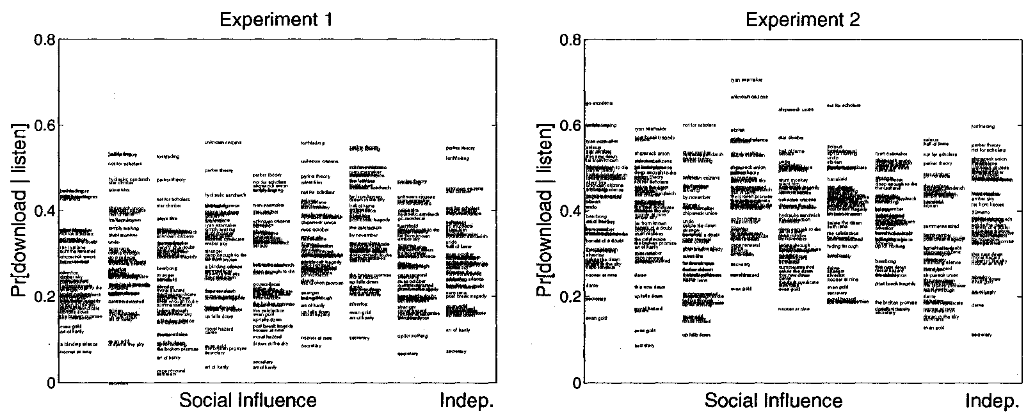


Figure B.8: Multi-thermometer plot of the probability of download given listen for each song in each world in experiment 2.



(a) Experiment 1

(b) Experiment 2

Figure B.9: Multi-thermometer plot of the probability of download given listen for each song in each world in experiments 1 and 2.

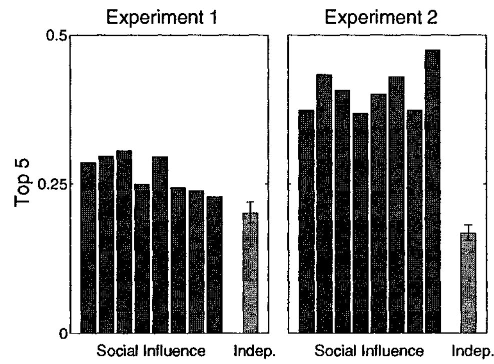


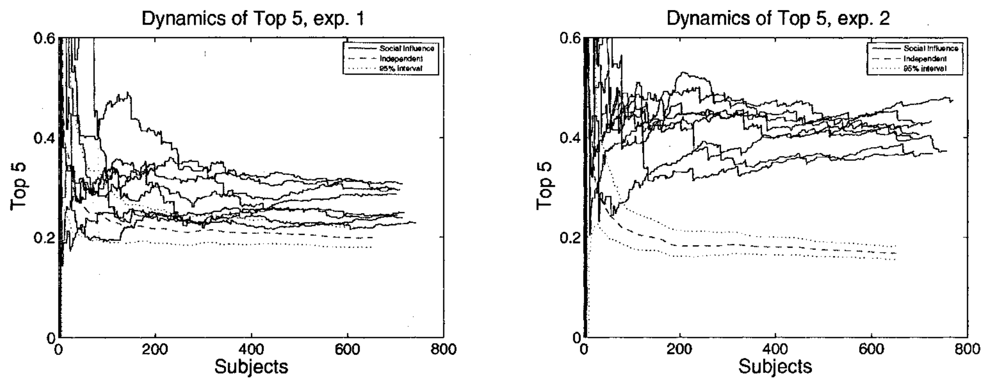
Figure B.10: Total market share of the top 5 songs in experiment 1 and 2.

B.2 Other measures of variability

As stated in the main text, there are many possible measures of inequality (Allison, 1978; Coulter, 1989). We chose to present the results in terms of the Gini coefficient, but we also checked to see if our results were robust to the specific measure of inequality used. To summarize our findings, the results are qualitatively unchanged for the three additional measures we investigated: market share of the top 5 songs, Herfindahl index, and Coefficient of variation. More details plots are available in figures B.10, B.11, B.12, and B.13 (Market share of top 5 songs); figures B.14, B.15, B.16, and B.17 (Herfindahl index); and figures B.18, B.19, B.20, and B.21 (Coefficient of variation).

B.3 Further analysis of listen choice data

The data in chapter 3 on which songs subjects chose to listen to (figures 3.11 and 3.18) were smoothed to improve presentation. The smoothing procedure was based on combining the actions from adjacent ranks and is best illustrated by an example. The probability of listening to the 3rd most popular song is the number of times a participant chose to listen to that song divided by the number of times that



(a) Experiment 1 (b) Experiment 2

Figure B.11: Dynamics of the total market share of the top 5 songs in experiments 1 and 2.

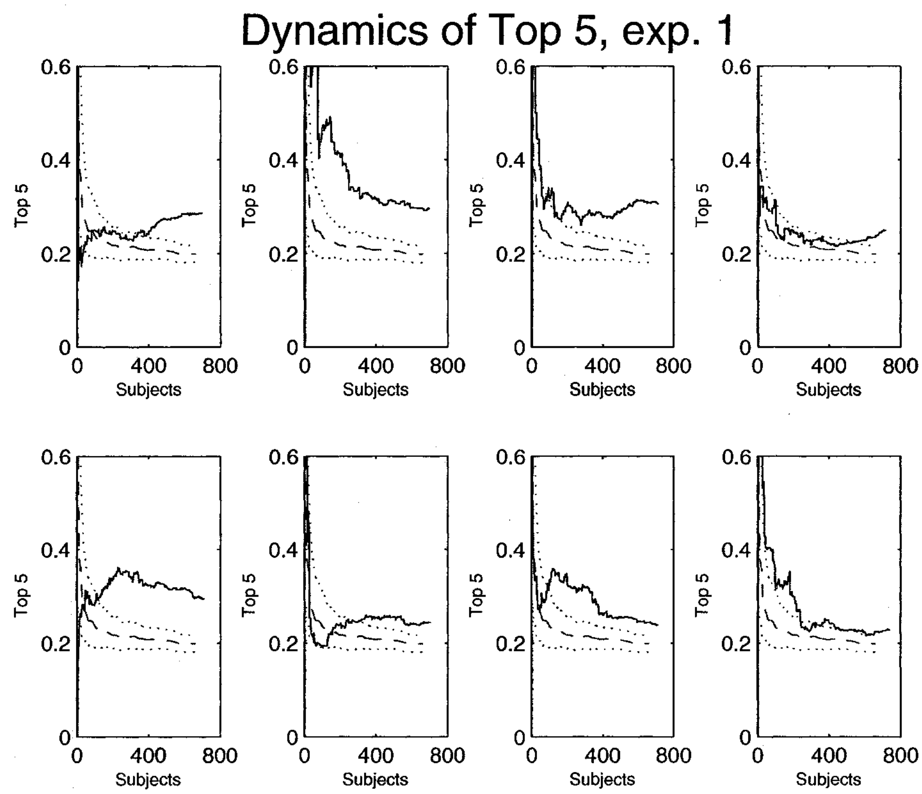


Figure B.12: Dynamics of the total market share of the top 5 songs in all eight worlds in experiment 1.

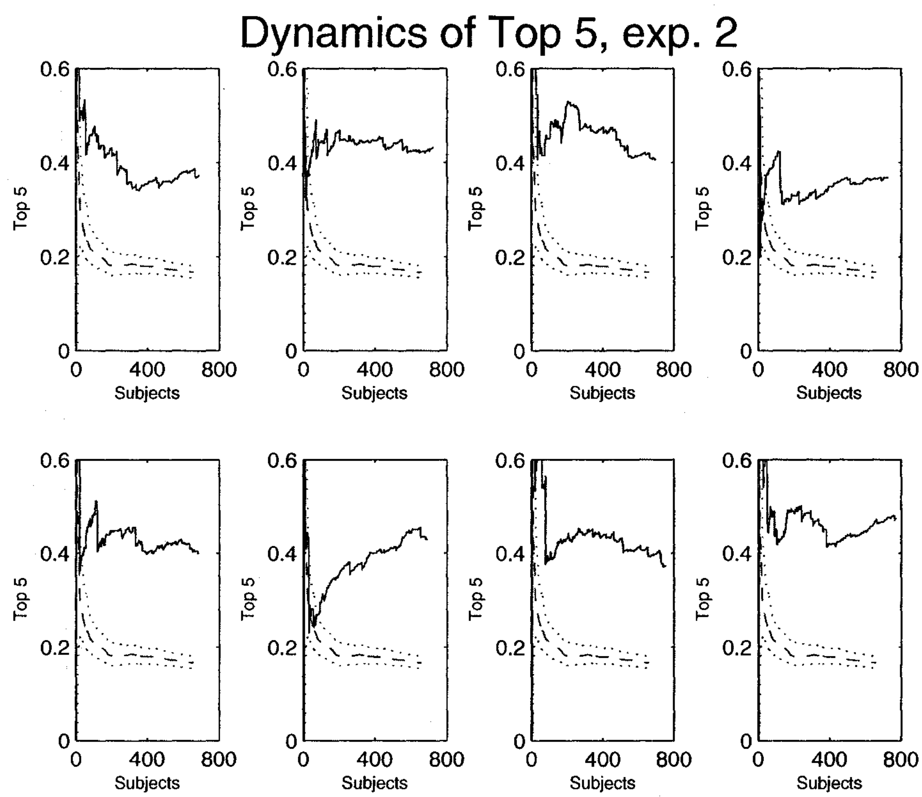


Figure B.13: Dynamics of the total market share of the top 5 songs in all eight worlds in experiment 2.

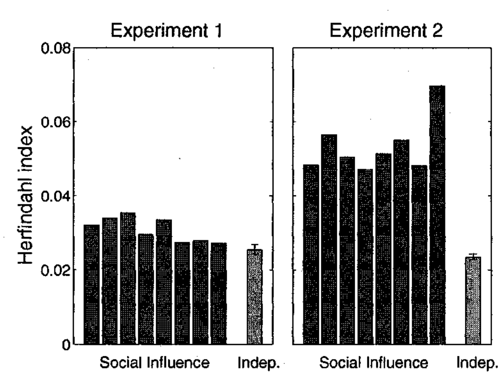
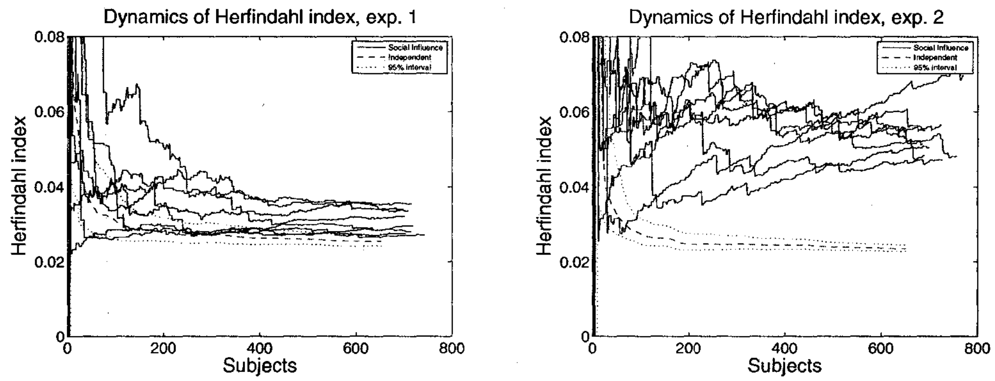


Figure B.14: Herfindahl index in experiment 1 and 2.



(a) Experiment 1

(b) Experiment 2

Figure B.15: Dynamics of the Herfindahl index in experiments 1 and 2.

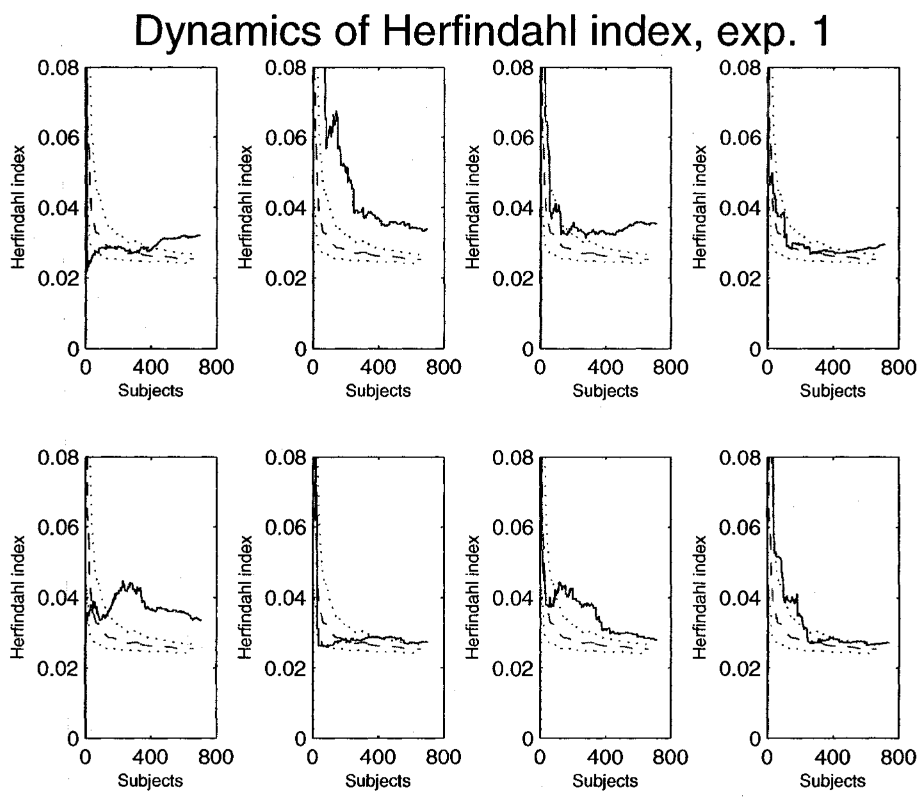


Figure B.16: Dynamics of the Herfindahl index in all eight worlds in experiment 1.

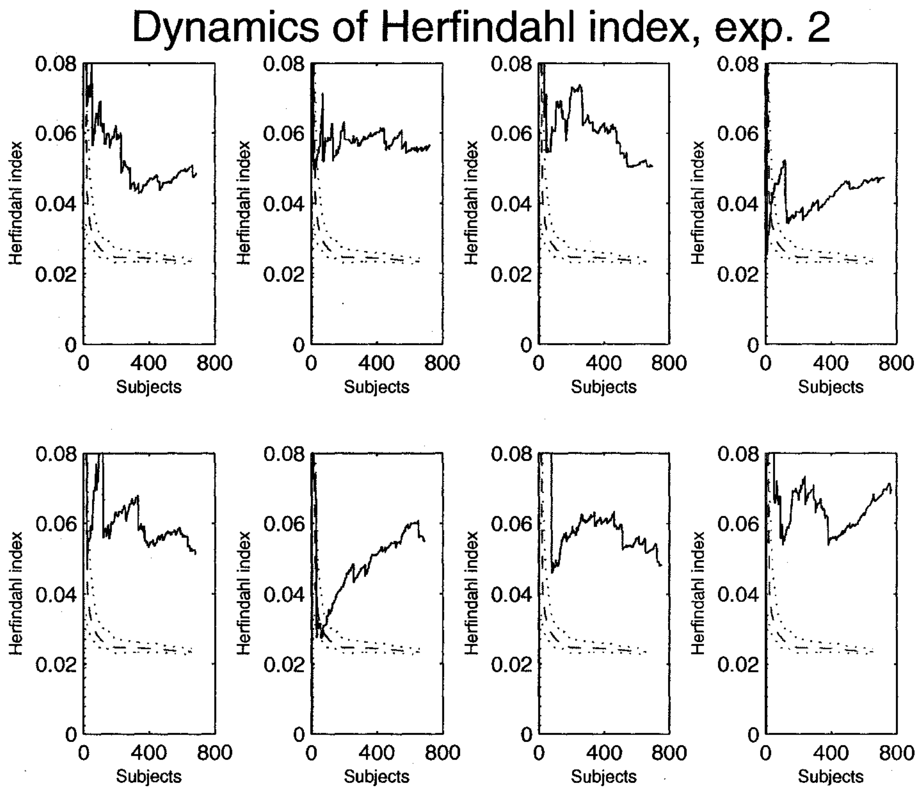


Figure B.17: Dynamics of the Herfindahl index in all eight worlds in experiment 2.

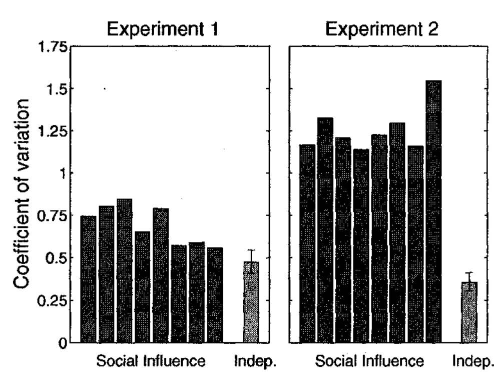
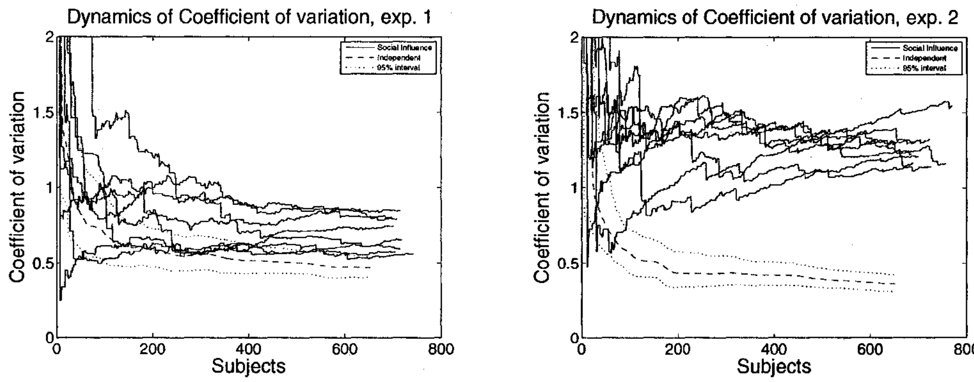


Figure B.18: Coefficient of variation in experiment 1 and 2.



(a) Experiment 1 (b) Experiment 2

Figure B.19: Dynamics of the coefficient of variation in experiments 1 and 2.

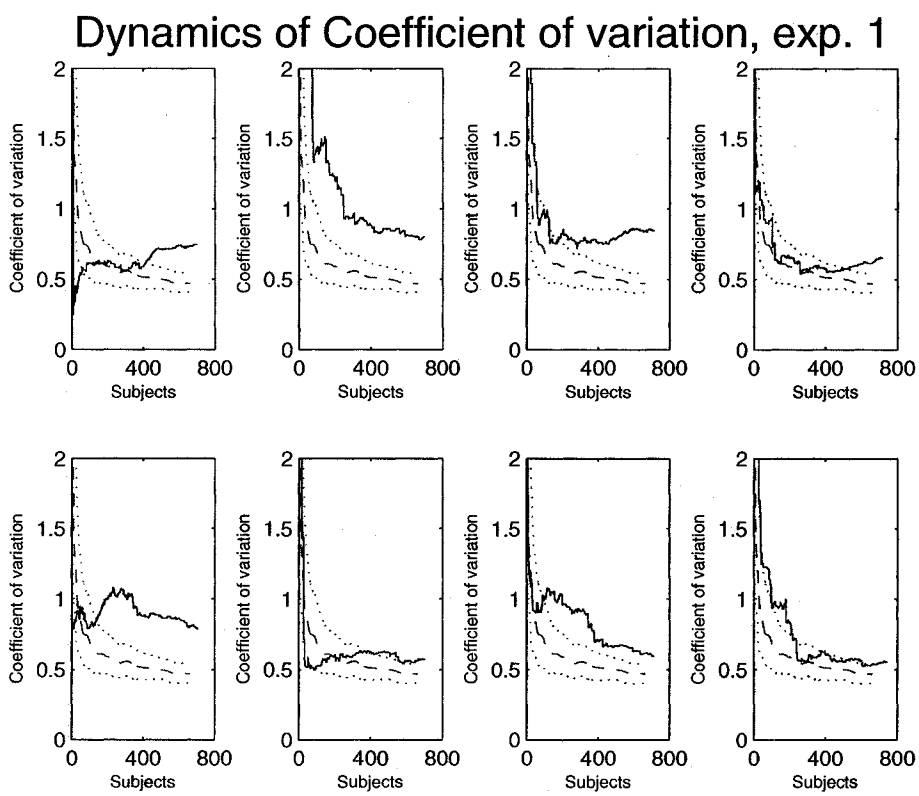


Figure B.20: Dynamics of the coefficient of variation in all eight worlds in experiment 1.

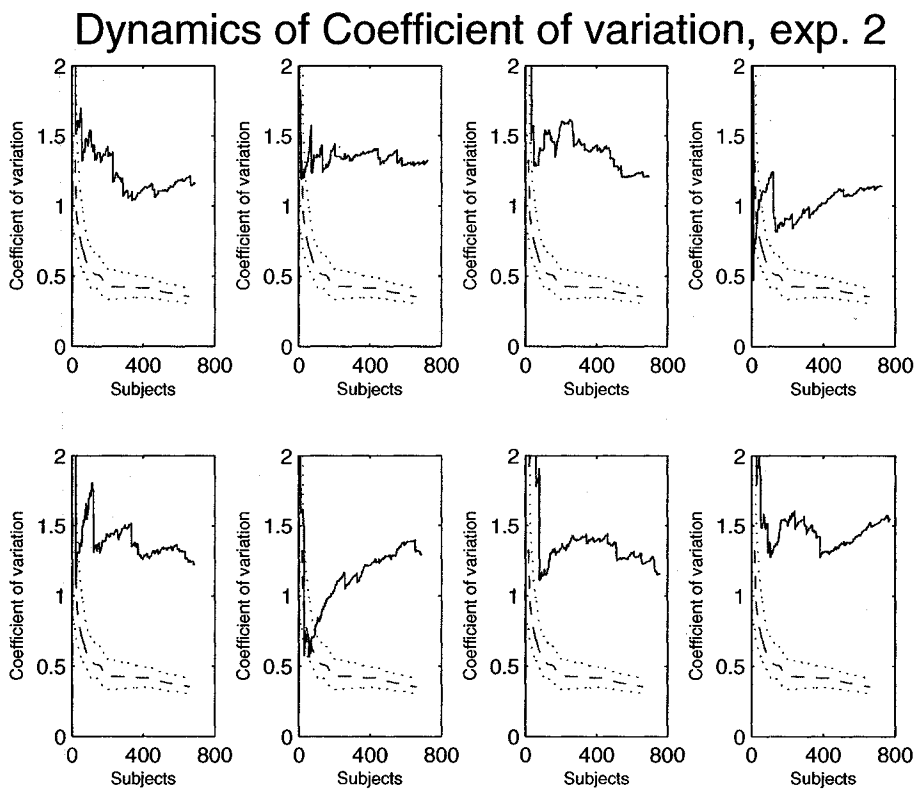


Figure B.21: Dynamics of the coefficient of variation in all eight worlds in experiment 2.

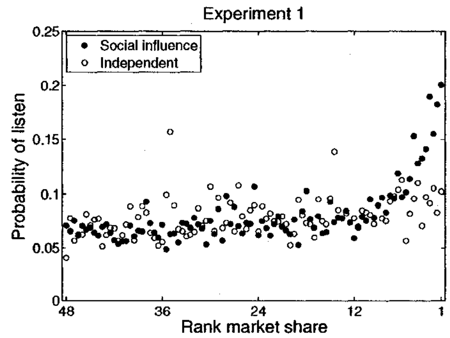
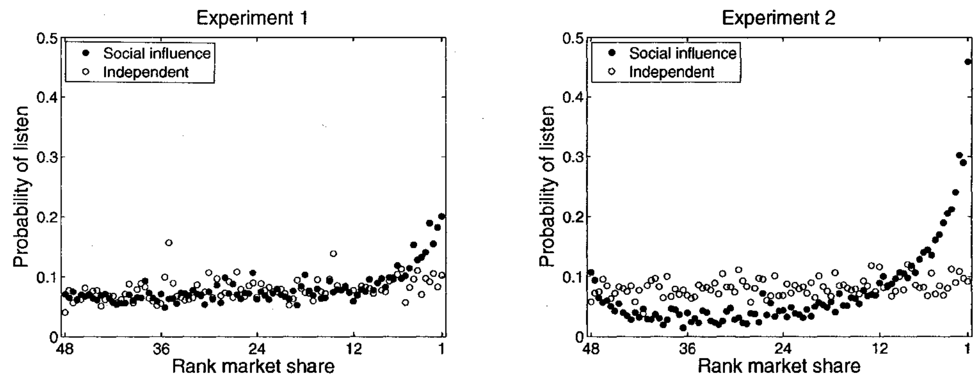


Figure B.22: Probability that a subject in each condition listened to a song of a given market rank in experiment 1 (unsmoothed). Smoothed results, which are qualitatively the same, are presented in figure 3.11.

there was a 3rd most popular songs (recall that because of ties there was not always a 3rd most popular song). In the smoothing procedure, we combined values so that the numerator became the number of times a participant chose to listen to the 3rd ranked song or the 2.5th or 3.5th ranked songs (following common practice (Kendall and Gibbons, 1990), we assigned a rank of 2.5 if two songs were tied for 2nd and 3rd and 3.5 if two songs were tied for 3rd and 4th). The denominator then became the number of times a song was ranked 3rd, or 2.5th or 3.5th. Thus, each position was averaged with its neighbors on either side. Note that the first and last ranked songs only had one neighbor, and so were only averaged over this one neighbor.

This smoothing procedure, while making the results easier to see, did not affect our substantive conclusions. The same results without smoothing are presented in figures B.22 and B.23. Further, figures B.24 and B.25 plot the unsmoothed data for each world separately.



(a) Experiment 1

(b) Experiment 2

Figure B.23: Probability that a subject in each condition listened to a song of a given market rank in experiment 1 and 2 (unsmoothed). Subjects' listening decisions were more strongly influenced by market rank in experiment 2 than experiment 1. Smoothed results, which are qualitatively the same, are presented in figure 3.18.

B.4 Download choice behavior and attachment kernels

We can also preform a similar analysis for the probability that a subject will download a song of a given rank. This so called attachment kernel is of interest in many models of cumulative advantage (Krapivsky and Redner, 2001). For the sake of completeness, results are presented in figures B.26, B.27, B.28,and B.29 and are unsmoothed.

B.5 A note on dynamics figures

When showing the dynamic values for the independent condition, the final value in the dynamic plot (for example, figure 3.12) is not the same as the value presented in the histogram plot (for example, 3.13). The reason for the difference is that the subsampled worlds can have different numbers of subjects. The dynamic plot is the length of the smallest subsampled world, but the value in the histogram is

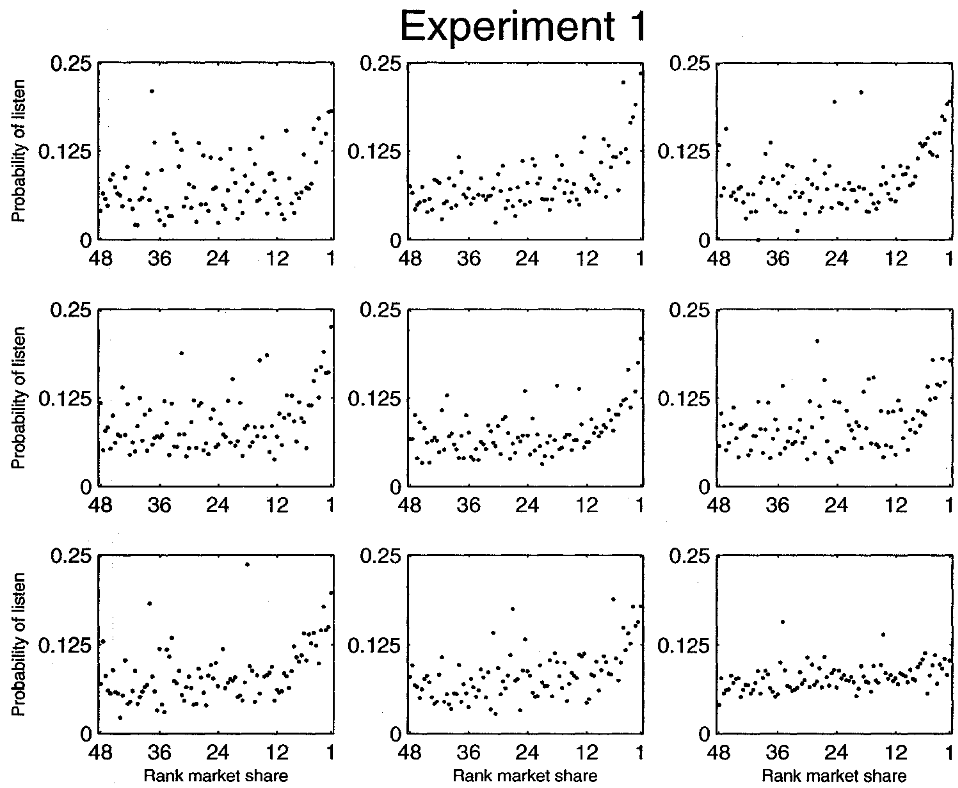


Figure B.24: Probability that a subject in each world listened to a song of a given market rank in experiment 1. These plots are not smoothed.

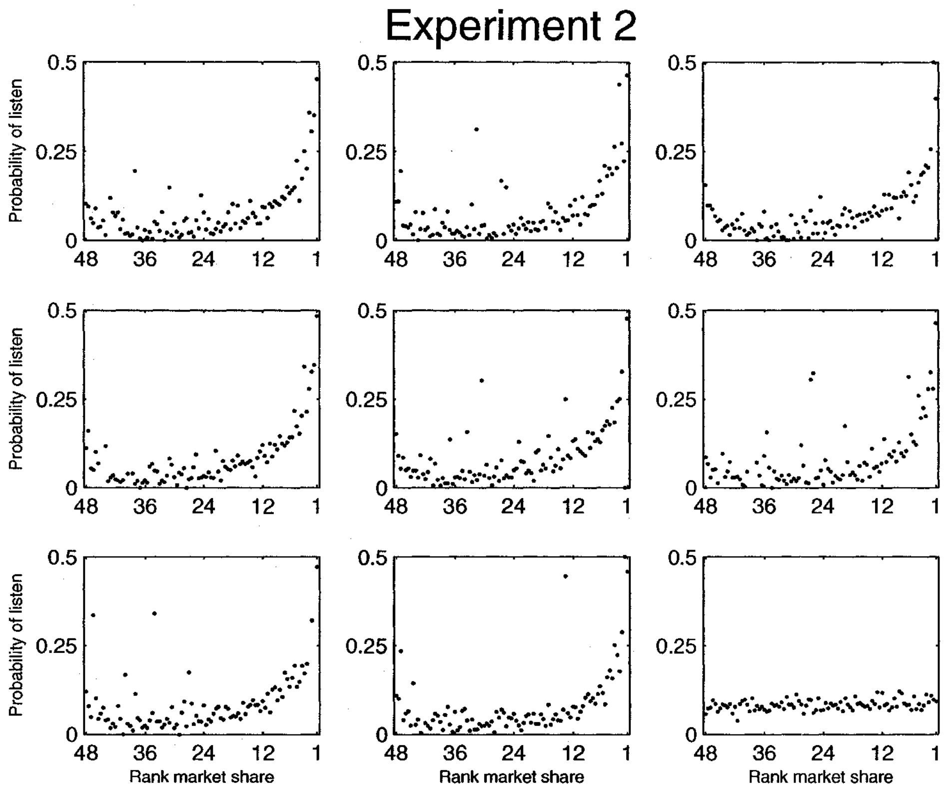


Figure B.25: Probability that a subject in each world listened to a song of a given market rank in experiment 2. These plots are not smoothed.

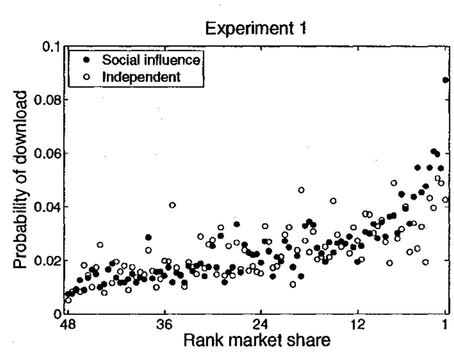
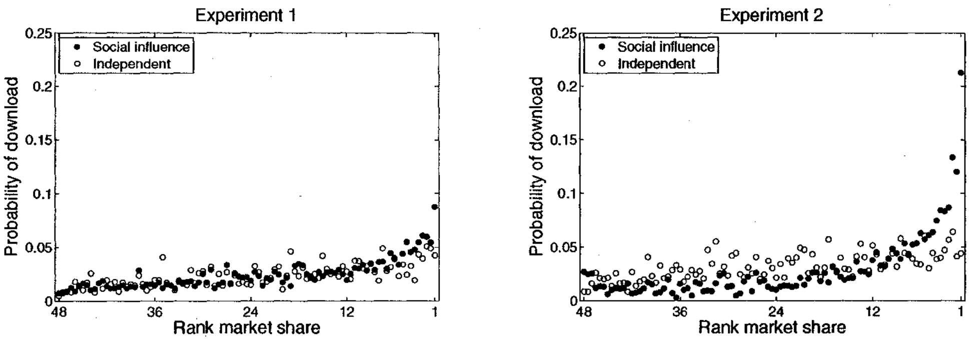


Figure B.26: Probability that a subject in each condition downloaded a song of a given market rank in experiment 1. This plot is not smoothed.



(a) Experiment 1 (b) Experiment 2

Figure B.27: Probability that a subject in each condition downloaded a song of a given market rank in experiment 1 and 2. This plot is not smoothed.

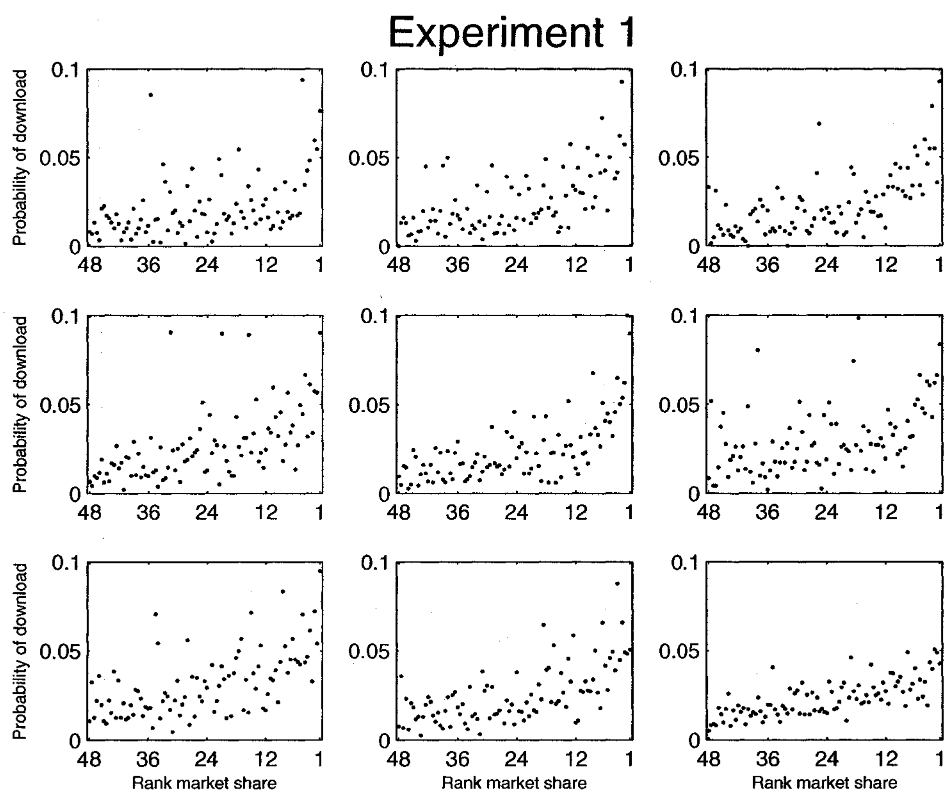


Figure B.28: Probability that a subject in each world downloaded a song of a given market rank in experiment 1. These plots are not smoothed.

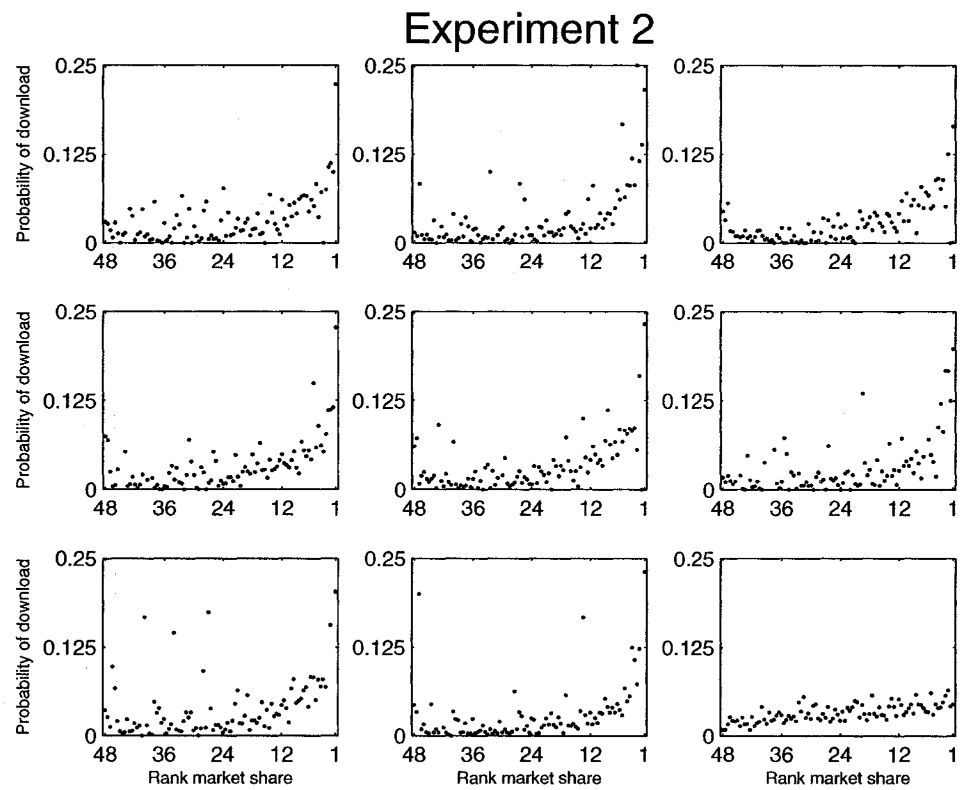


Figure B.29: Probability that a subject in each world downloaded a song of a given market rank in experiment 2. These plots are not smoothed.

based on the the ending value in each subsampled world. This discrepancy is small and not a cause for concern, but it is worth noting.

Appendix C

Appendix to chapter 4

C.1 Other presentations of success outcomes

Table C.1 reports the top 10 songs in each world in experiment 3. This same information is presented in a multiple-thermometer plot in figure C.1, and figure C.2 presents this multiple thermometer plot for experiment 3 next to equivalent plot for experiment 2. Figure C.3 presents a multiple-thermometer plot of the mean rating of each song in each world, and figure C.4 presents the multiple thermometer plot for experiment 3 next to equivalent plot for experiment 2. Finally, figure C.5 presents a multiple-thermometer plot of the probability of download given listen (i.e., batting average) fom each song in each world, and figure C.6 presents the multiple thermometer plot for experiment 3 next to equivalent plot for experiment 2.

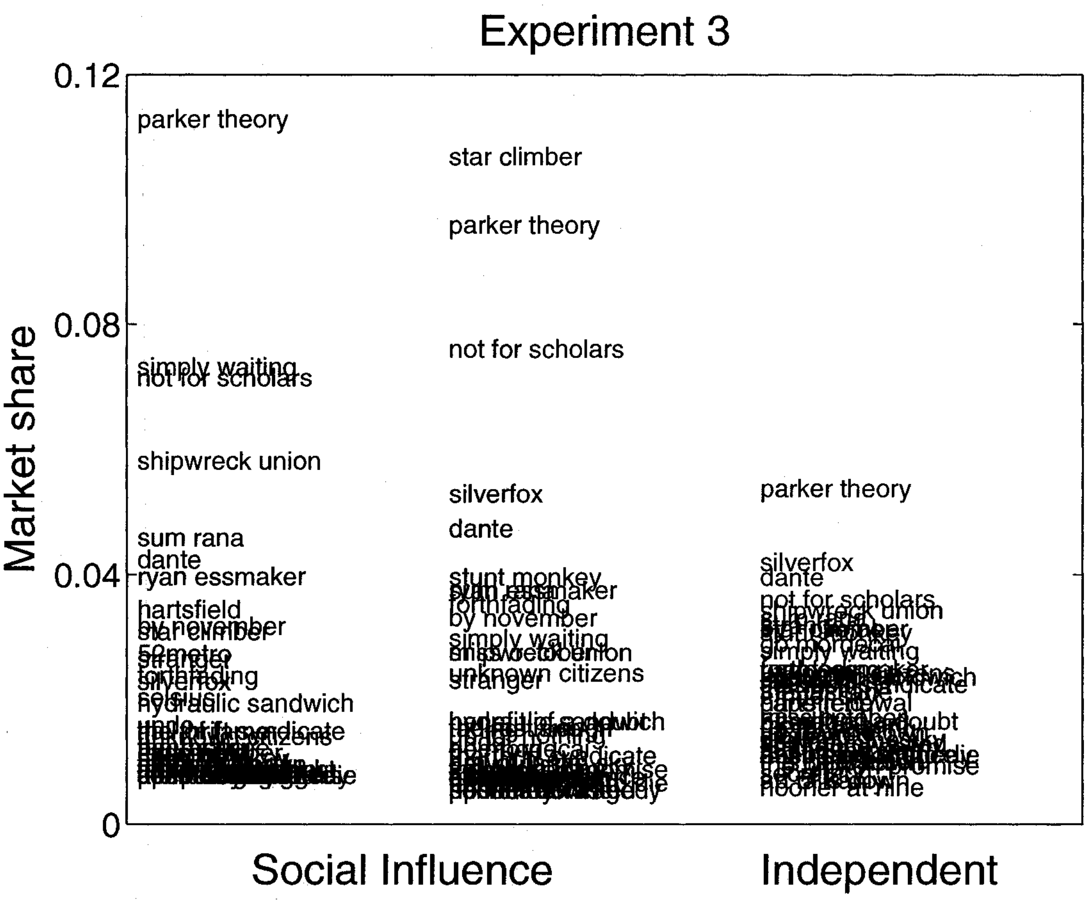
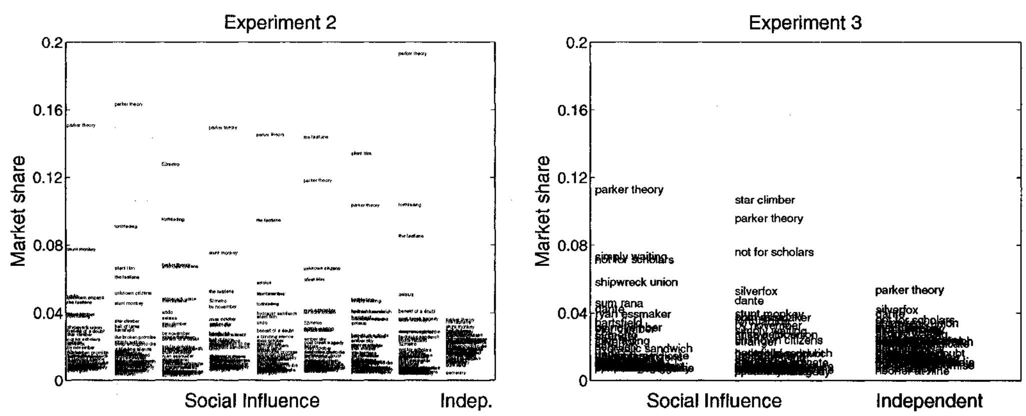


Figure C.1: Multi-thermometer plots of market share in experiment 3.

Experiment 3

	Social Influence		
Rank	World 1	World 2	Independent
1	parker theory	star climber	parker theory
2	simply waiting	parker theory	silverfox
3	not for scholars	not for scholars	dante
4	shipwreck union	silverfox	not for scholars
5	sum rana	dante	shipwreck union
6	dante	stunt monkey	sum rana
7	ryan essmaker	sum rana	stranger
8	hartsfield	ryan essmaker	by november
9	by november	forthfading	star climber
10	star climber	by november	stunt monkey

Table C.1: Top 10 songs in each world in experiment 3. In the two social influence worlds there were two different first place songs.



(a) Experiment 2 (b) Experiment 3

Figure C.2: Multi-thermometer plots of market share in experiment 2 and 3.

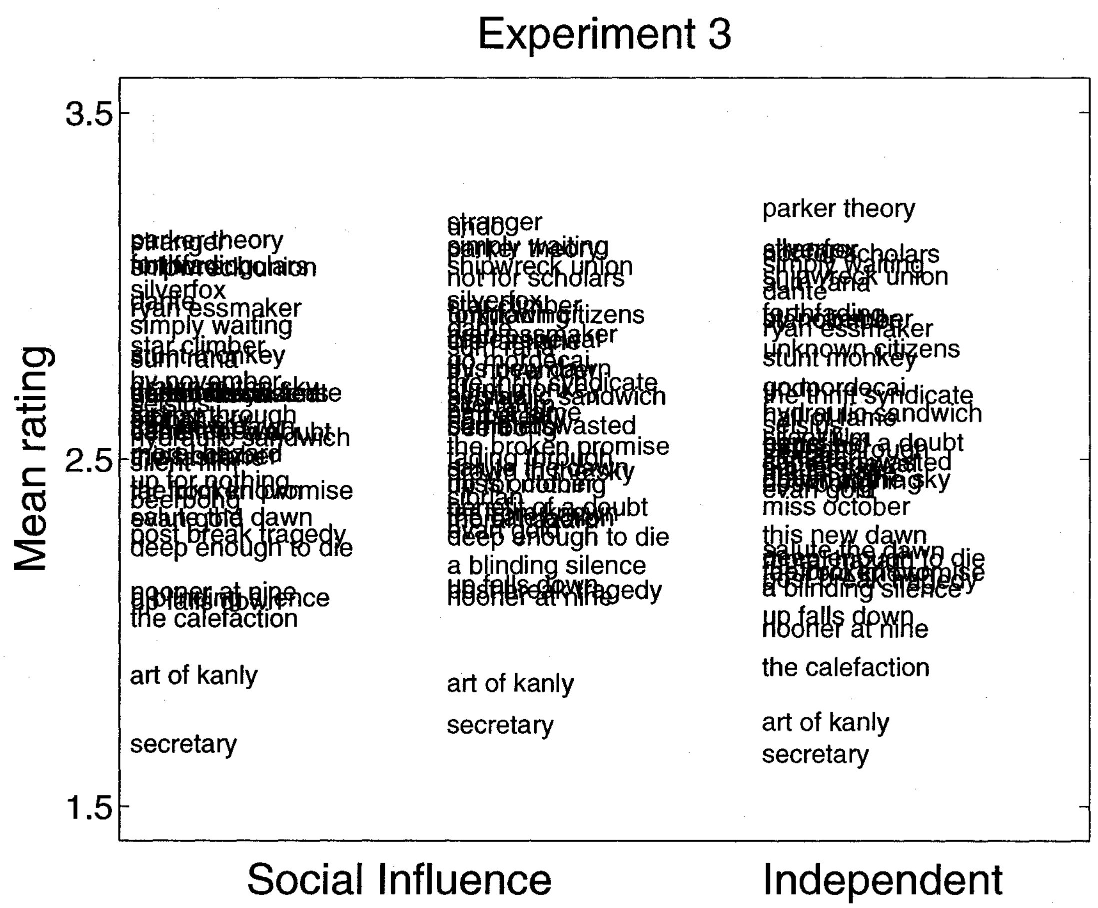
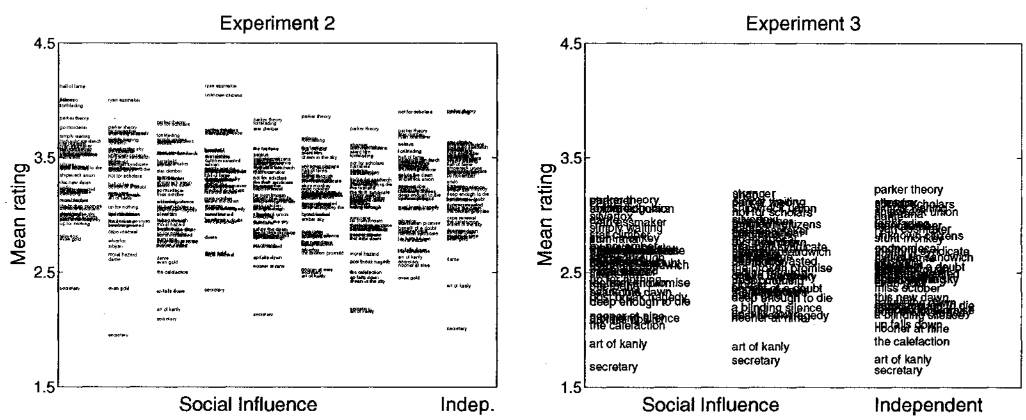


Figure C.3: Multi-thermometer plot of mean rating of each song in each world in experiment 3.



(a) Experiment 2 (b) Experiment 3

Figure C.4: Multi-thermometer plot of mean rating of each song in each world in experiments 2 and 3.

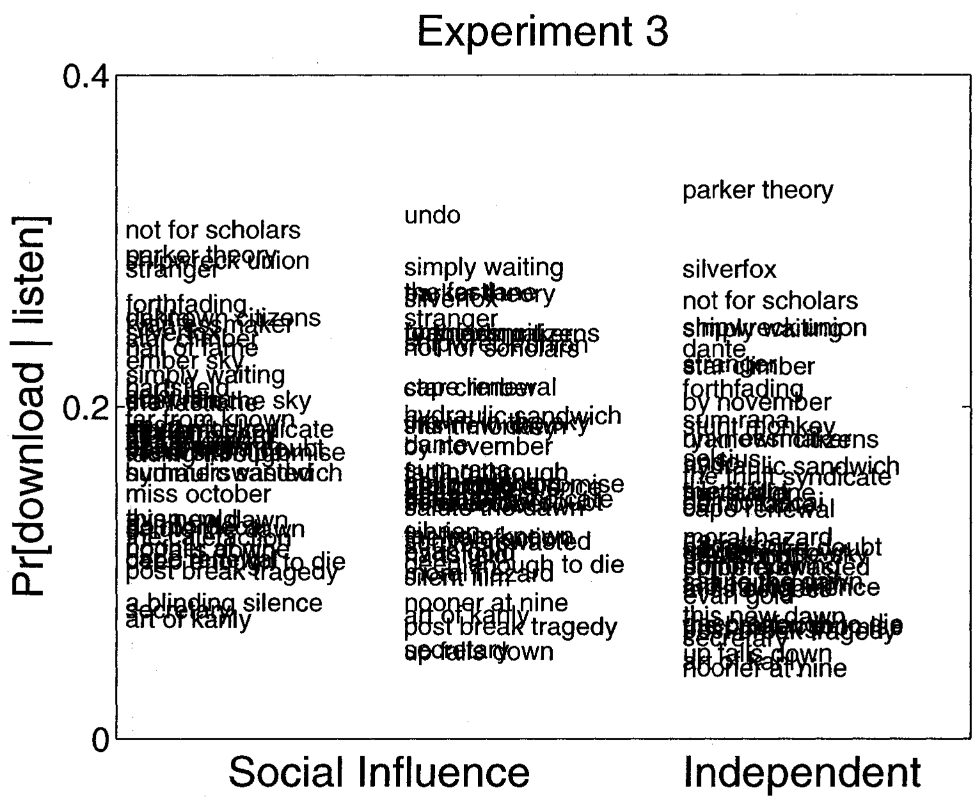
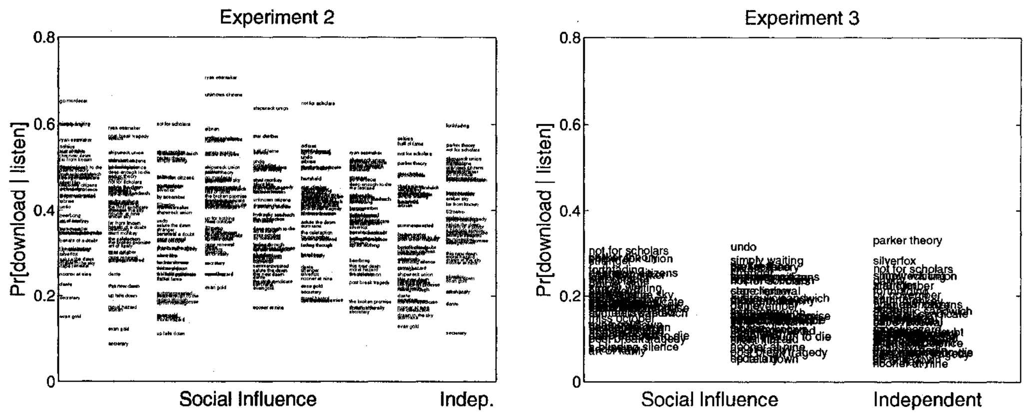


Figure C.5: Multi-thermometer plot of the probability of download given listens for each song in each world in experiment 3.



(a) Experiment 2 (b) Experiment 3

Figure C.6: Multi-thermometer plot of the probability of download given listen for each song in each world in experiments 2 and 3.

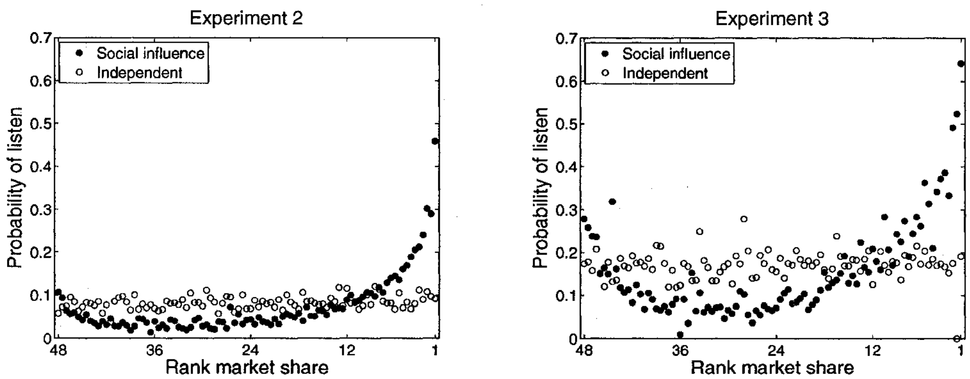
C.2 Email to electronic small-world subjects

Hello,

Thank you for your previous participation in the Small World Research Project testing the idea that any two people in the world can be connected by 'six degrees of separation.'

Now we have another experiment that we think you might like. It's all about how people evaluate music. If you visit the site, you'll get a chance to listen to, and download, music by cool new artists. Of course, to download the music is 100% free and 100% legal. Check it out, <web address here>

Thanks,
The Small World team
http://smallworld.columbia.edu



(a) Experiment 2 (b) Experiment 3

Figure C.7: Probability that a subject in each condition listened to a song of a given market rank. Smoothed results are presented in figure 4.11.

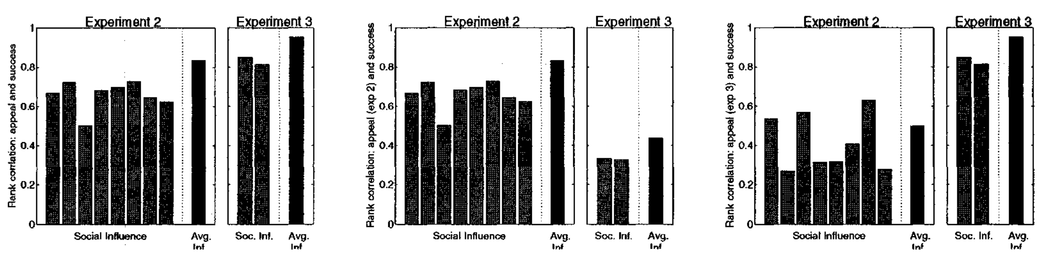
`If you do not wish to receive emails from us in the future, please click here <url here>.`

C.3 Unsmoothed listen choice plots

In the main text, we presented smoothed listen choice plots (figure 4.11). Figure C.7 presents the unsmoothed plots which are qualitatively the same.

C.4 Predicting market rank as measured by rank-correlation

In section 4.5.3 we calculated the probability of correctly predicting the ordering of success of two songs in one outcome given their ordering in some other outcome. A related measure is Spearman's rank correlation (Kruskal, 1958). Figure C.8 plots the rank correlation between appeal and success within and across experiments.



(a) Correct appeal (b) Exp 2 appeal (c) Exp 3 appeal

Figure C.8: The left figure presents the rank correlation between appeal from experiment 2 and success in experiment 2 as well as the rank correlation between appeal from experiment 3 and success from experiment 3. The center figure presents the rank correlation between appeal from experiment 2 and success in experiment 2 and 3. The right figure presents the rank correlation between appeal from experiment 3 and success in experiment 2 and 3.

Appendix D

Appendix to chapter 5

D.1 Band popularity check

As in the previous experiments, we checked to ensure that the bands used in the experiment were unknown to most participants. 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).[1] Table D.1 reports results from the 2,221 subjects from experiment 3 who make up the set-up period for experiment 4 and the 9,996 subjects in experiment 4 itself. The table shows that some subjects reported being familiar with the three potential bands, but these recognition rates were no higher than for the imaginary band.[2] Further, the extremely different results observed for the band U2 suggest that respondents were actually reading the question and not simply reporting "don't

[1]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 (Goldstein and Gigerenzer, 2002).

[2]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.

	How familiar are you with the following bands?		
	Don't know it at all (% of subjects)	Heard of it (% of subjects)	Know it pretty well (% of subjects)
Real Bands			
Guys on Couch	94.3	4.9	0.8
Grover Dill	94.4	4.9	0.7
U2	2.4	18.2	79.4
Remnant Soldier	90.0	8.9	1.2
Fake Band			
Peter on Fire	92.0	6.9	1.1

Table D.1: Comparing the popularity of the three potential bands from our sample (Guys on Couch, Grover Dill, Remnant Soldier) to a fake band (Peter on Fire) and a popular real band (U2). Subjects reported being about as familiar with an fake band as three potential bands from our sample. Totals may not sum to 100 because of rounding. Results are based on 2,211 subjects before the inversion and 9,996 subjects after the inversion.

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.

D.2 Other presentations of success outcomes

Tables D.2 and D.3 report the top 10 songs in each world in experiment 4 either including or excluding the initial conditions. This same information is presented in a multiple-thermometer plot in figures D.1 and D.2, and figure D.3 presents these two plots side-by-side to allow for easy comparison. Figure D.4 present a multiple-thermometer plot of the mean rating of each song in each world and figure D.5 presents a multiple-thermometer plot of the probability of download given listen (i.e., batting aveage) for each song in each world.

Experiment 4 (including initial conditions)

	Social Influence			
Rank	Unchanged	Inverted #1	Inverted #2	Independent
1	parker theory	post break tragedy	post break tragedy	parker theory
2	not for scholars	up for nothing	up for nothing	not for scholars
3	simply waiting	deep enough to die	parker theory	silverfox
4	shipwreck union	this new dawn	deep enough to die	dante
5	star climber	go mordecai	this new dawn	stunt monkey
6	sum rana	star climber	go mordecai	shipwreck union
7	dante	parker theory	sum rana	ember sky
8	ryan essmaker	not for scholars	cape renewal	stranger
9	hartsfield	up falls down	dante	ryan essmaker
10	stranger	a blinding silence	up falls down	star climber

Table D.2: Top 10 songs in each world in experiment 4 (including initial conditions).

Experiment 4 (excluding initial conditions)

	Social Influence			
Rank	Unchanged	Inverted #1	Inverted #2	Independent
1	parker theory	post break tragedy	parker theory	parker theory
2	not for scholars	star climber	post break tragedy	not for scholars
3	shipwreck union	up for nothing	up for nothing	silverfox
4	simply waiting	parker theory	sum rana	dante
5	star climber	not for scholars	go mordecai	stunt monkey
6	sum rana	go mordecai	dante	shipwreck union
7	dante	deep enough to die	not for scholars	ember sky
8	ryan essmaker	shipwreck union	cape renewal	stranger
9	forthfading	this new dawn	this new dawn	ryan essmaker
10	stranger	dante	by november	star climber

Table D.3: Top 10 songs in each world in experiment 4 (excluding initial conditions).

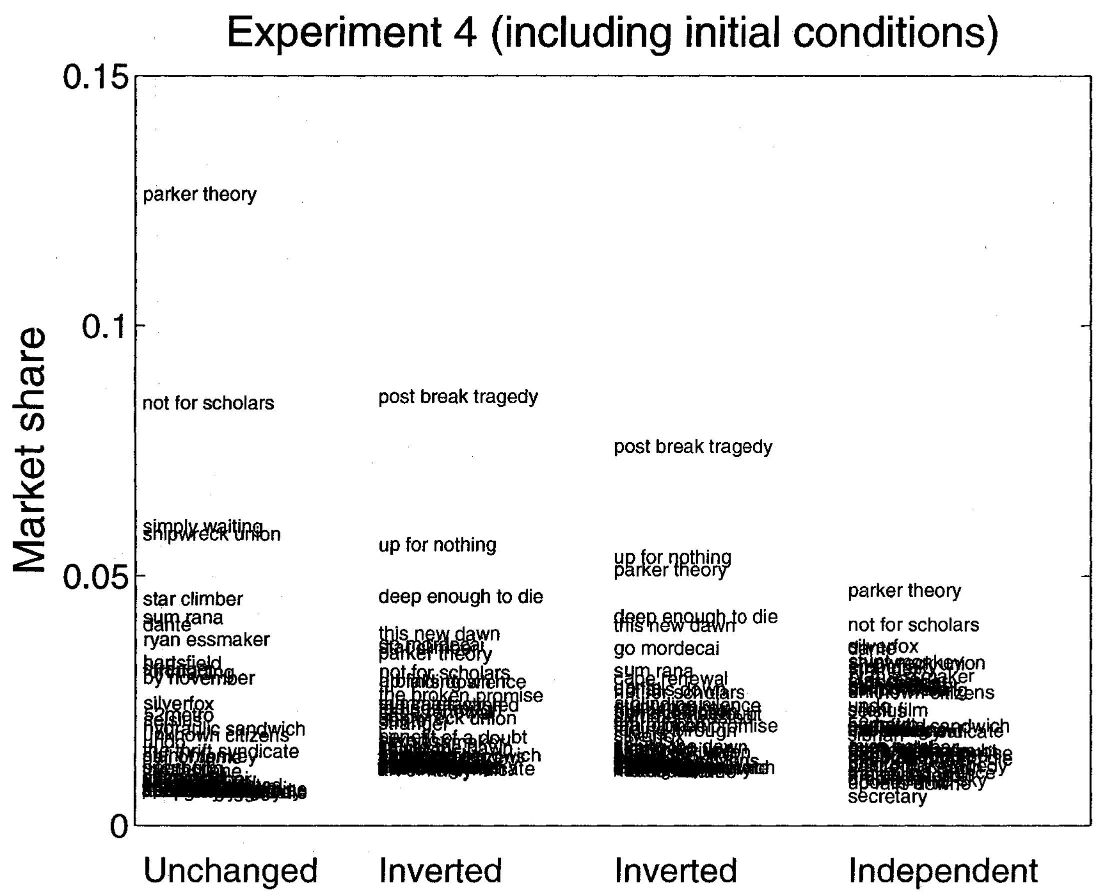


Figure D.1: Multi-thermometer plots of market share in experiment 4 (including initial conditions).

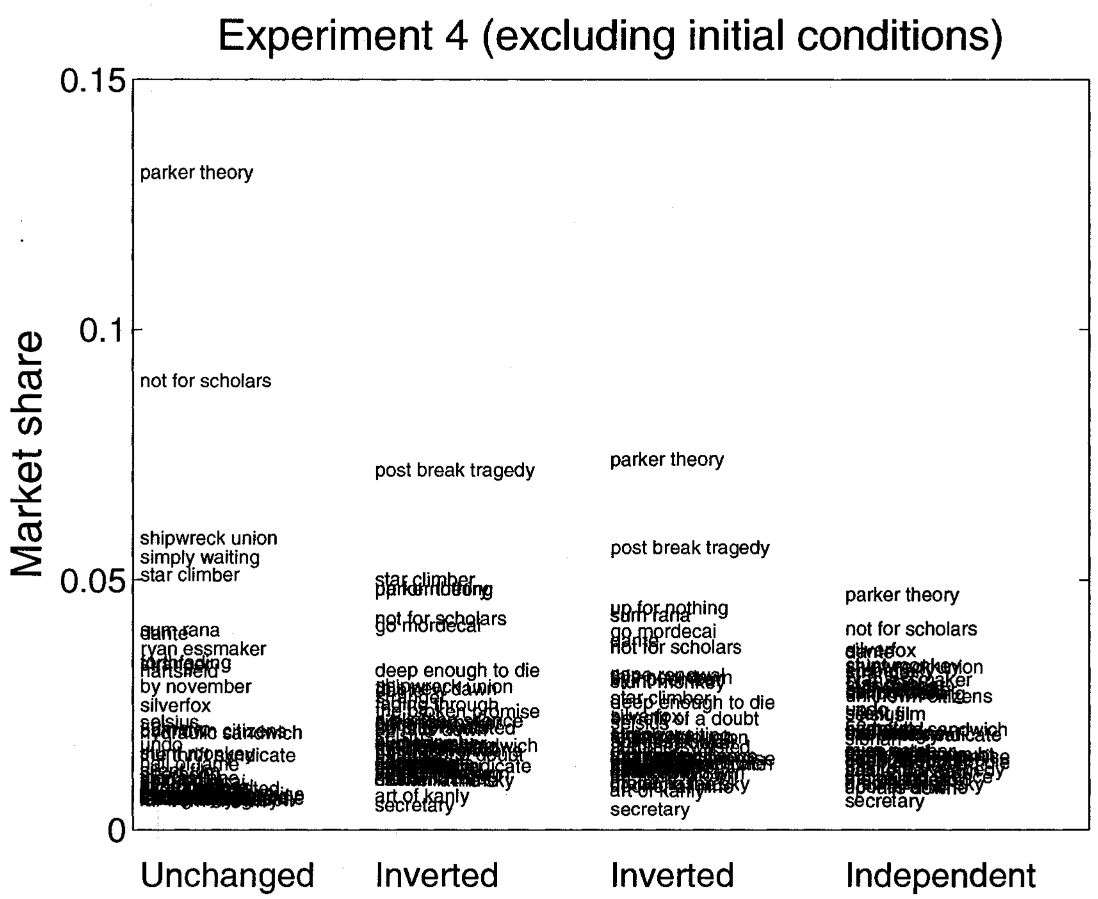
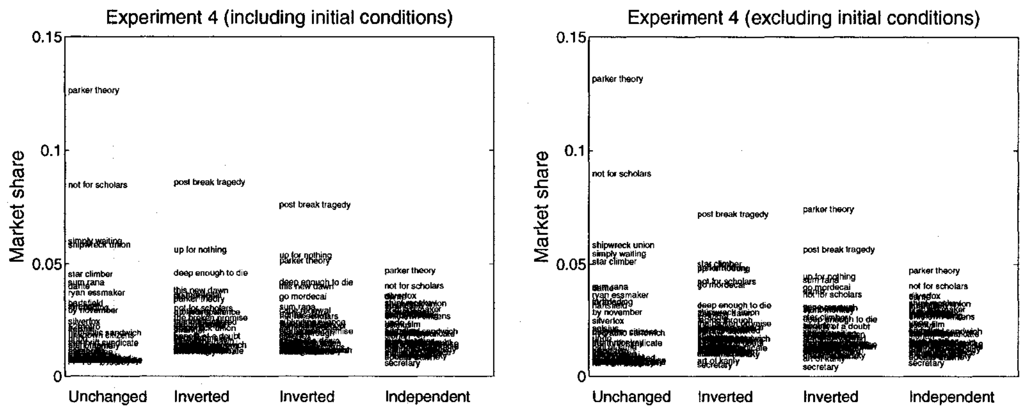


Figure D.2: Multi-thermometer plots of market share in experiment 4 (excluding initial conditions).



(a) Including initial conditions (b) Excluding initial conditions

Figure D.3: Multi-thermometer plots of market share in experiment 4 including and excluding initial condition.

D.3 More on the reduction in downloads in the inverted worlds

In the main text (section 5.5) we showed that participants in the inverted social influence worlds listened to and downloaded fewer songs than those in the unchanged world and we argue that this was because of the distortion of the social signal caused by the inversion. Very strong evidence for our claim would be if the subject behavior in the inverted worlds changed over time as the inversion was gradually undone (figure 5.25). Figure D.6 plots the number of listens for each subject in each world as the experiment progresses. There does not seem to be very strong evidence that subjects listens to more songs over time in the inverted worlds. Figure D.7 presents the same information for downloads. Again there is not a clear pattern.

Because the data a so noisy they may be masking a pattern. A simple check would be to compare the behavior of subjects in the first and second half of the experiment. In figure D.8(a) the grey bars show the average number of listens in the first half of each world and the black bars show the second half. In the inverted

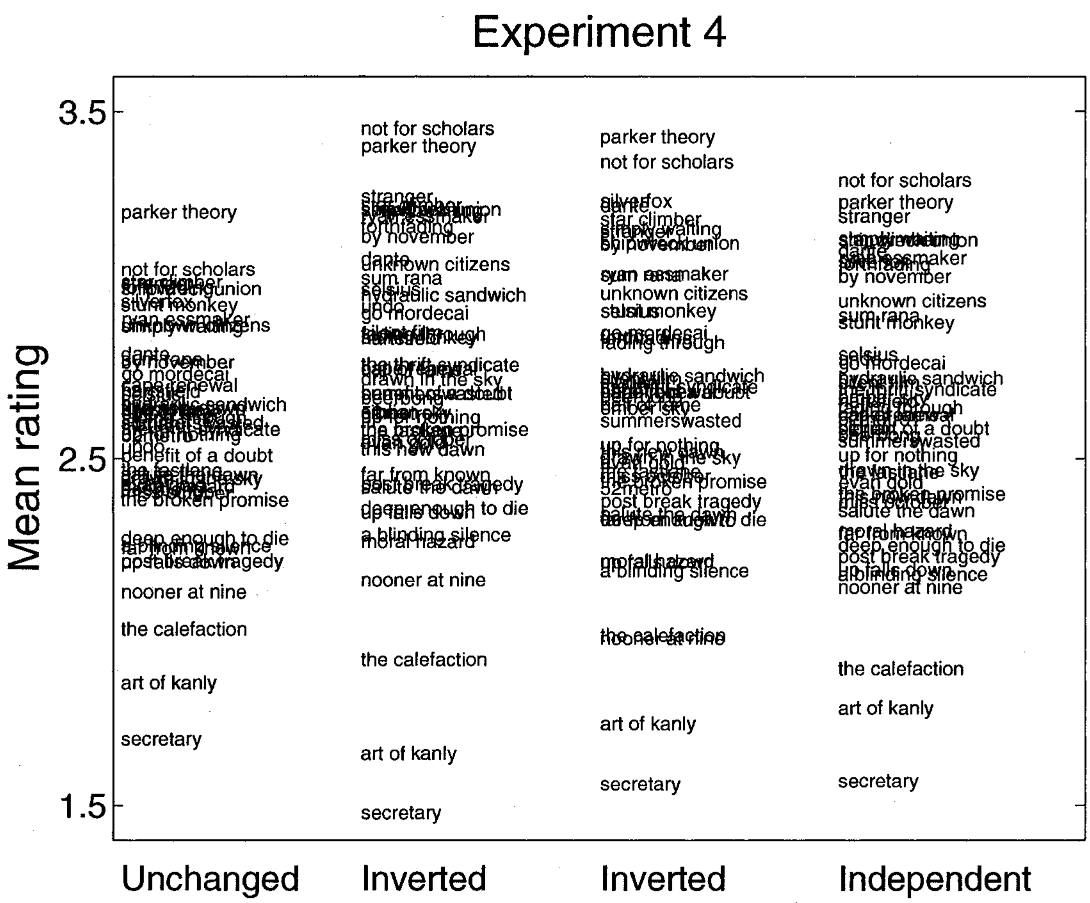


Figure D.4: Multi-thermometer plot of mean rating for each song in each world in experiment 4.

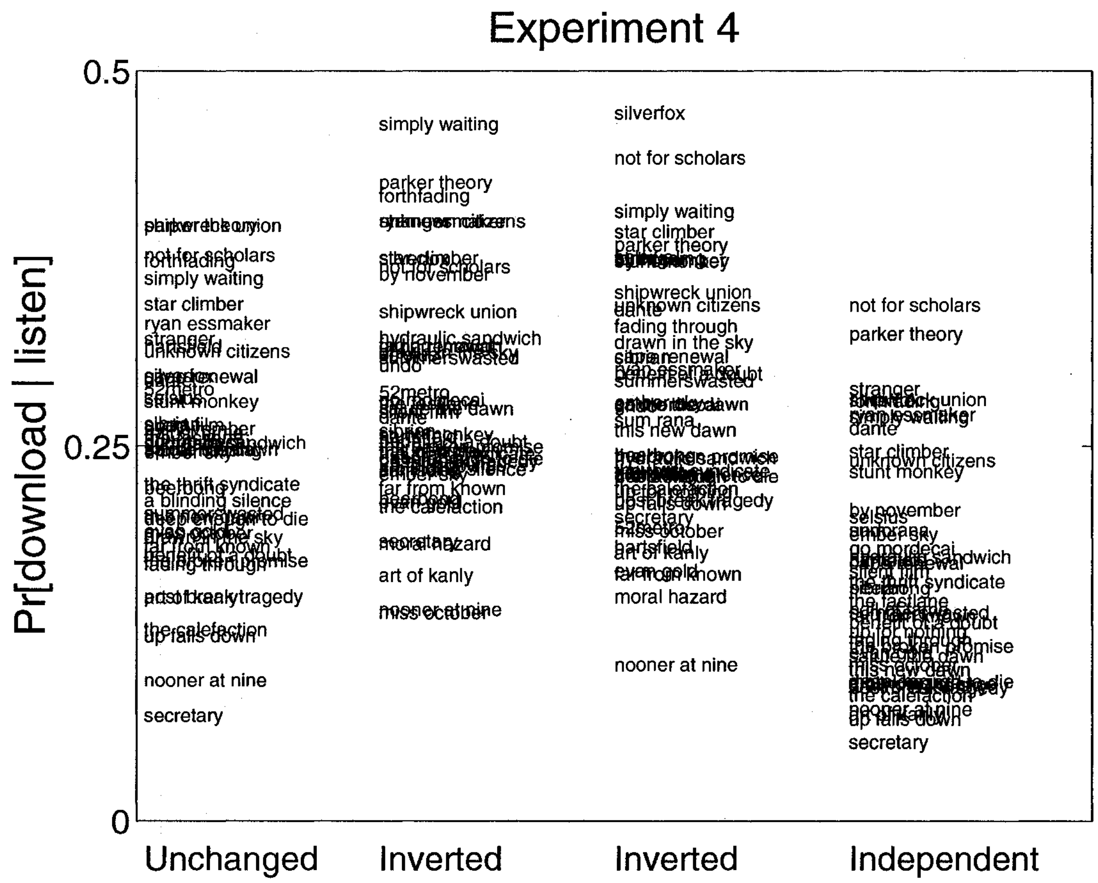
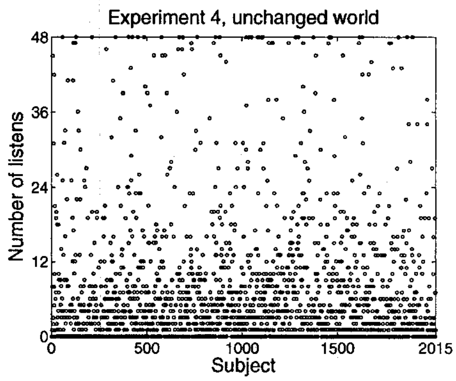
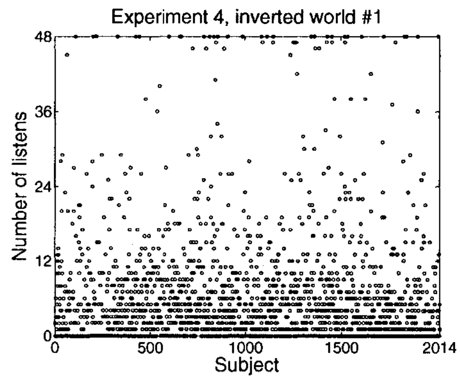


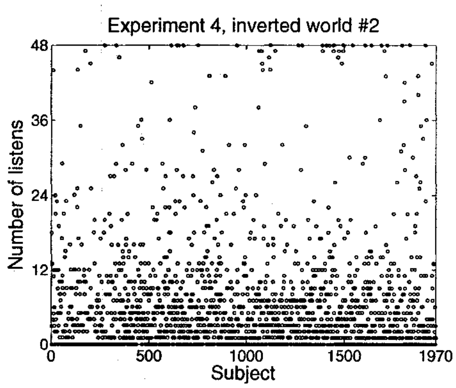
Figure D.5: Multi-thermometer plot of the probability of download given listen for each song in each world in experiment 4.



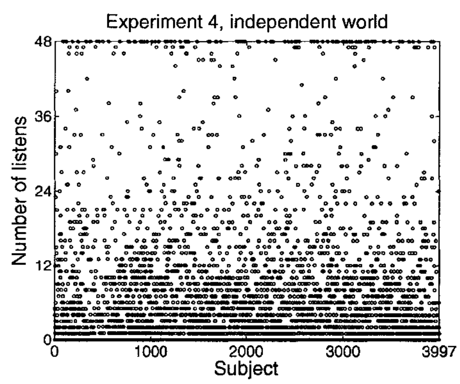
(a) Unchanged world



(b) Inverted world #1

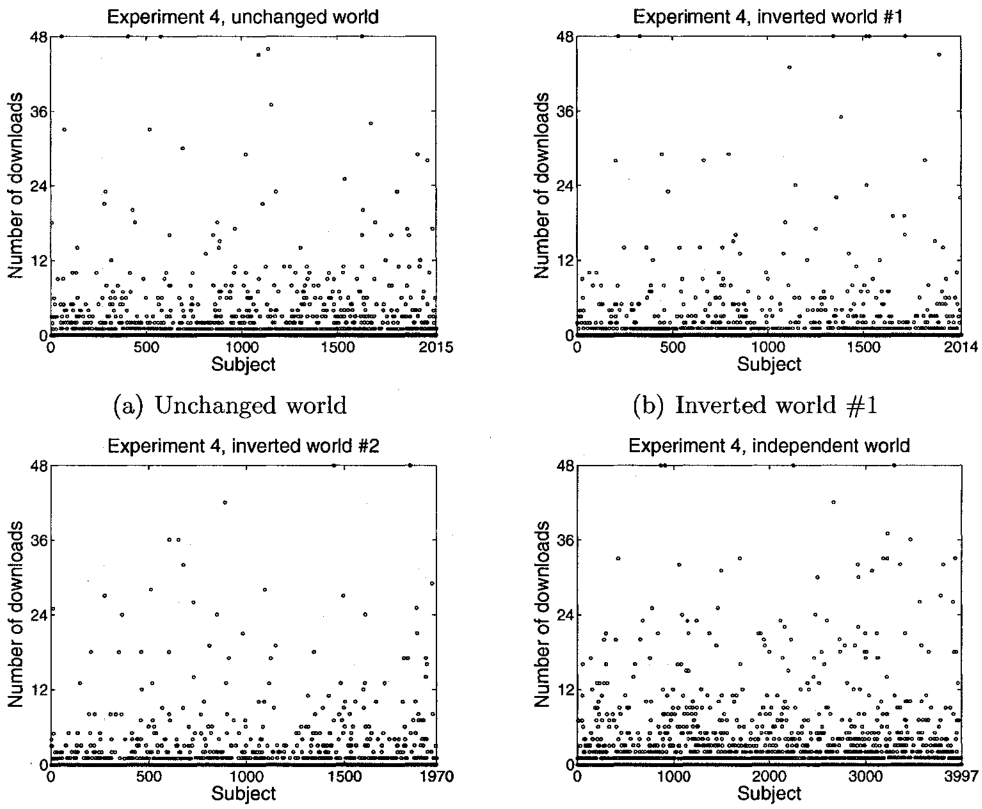


(c) Inverted world #2



(d) Independent world

Figure D.6: Number of listens for each subject in each world in experiment 4. It is not immediately clear that the number of listens per subject increases over time in the inverted worlds as they gradually return to their original order.



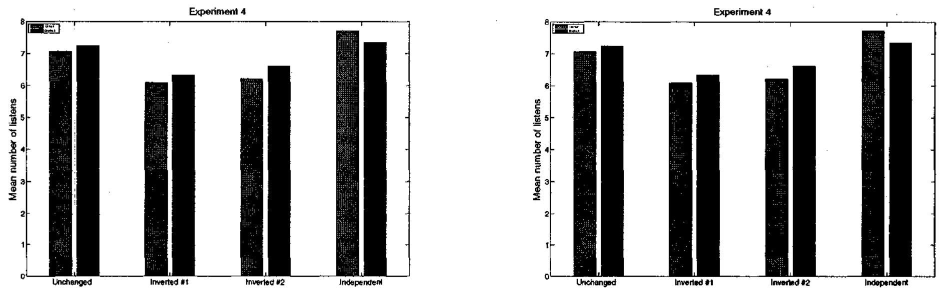
(a) Unchanged world

(b) Inverted world #1

(c) Inverted world #2

(d) Independent world

Figure D.7: Number of downloads for each subject in each world in experiment 4. It is not immediately clear that the number of downloads per subject increases over time in the inverted worlds as they gradually return to their original order.



(a) Listens (b) Downloads

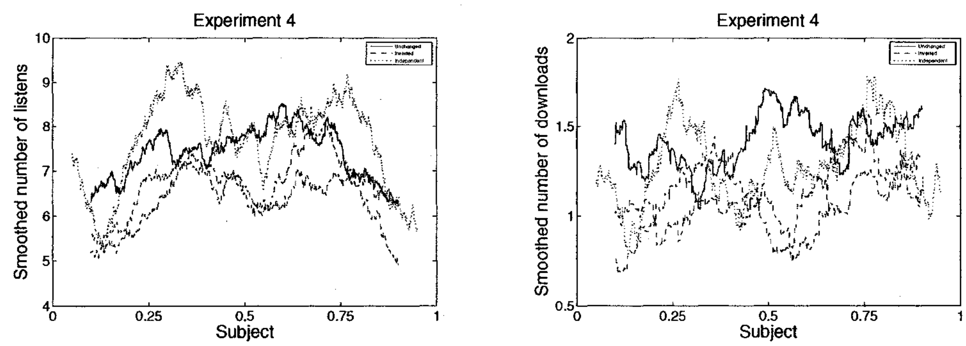
Figure D.8: Subject listen and download behavior in the first and second half of the four worlds in experiment 4. Subjects in the inverted worlds listened to and downloaded more in the second half (black bars) of the experiment than the first half (grey bars), but so did subjects in the unchanged world.

worlds, subjects did listen to more songs in the second half of the experiment, but the same pattern can be observed in the unchanged world. Similar patterns are observed in figure D.8(b) which plots subject download behavior.

Perhaps splitting each world in half is too coarse a check. To get a more refined notion of how behavior changed over time we also smoothed the data with a moving average smoother that for each subject averaged of the previous 200 subjects and the next 200 subjects. Thus each smoothed data point represents the average of 401 data points.[3] Figure D.6 shows that in all four worlds the number of listens increased and then decreased, but the pattern seems roughly the same for the inverted worlds as the unchanged world.[4] Figure D.6, which presented smoothed download behavior, also does not seem to show increased downloads in the inverted worlds (compared to

[3]There are certainly more sophisticated smoothing techniques, but for our purposes this simple one should be sufficient.

[4]The x-axis for this figure is subject number relative to the total number of subjects in that world. This way the independent world with $\sim$4,000 subjects can be plotted on the same axis as the social influence worlds with $\sim$2,000 subjects. Also, this difference in the number of subjects explains why the line for the independent world is longer than those of the social influence worlds; because the smoothing window size is a constant, this represents a smaller proportion of subjects in the independent world.



(a) Smoothed listens

(b) Smoothed downloads

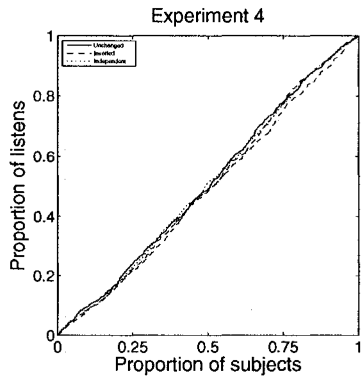
Figure D.9: Smoothed number of listens and downloads for each subject in experiment 4. It seems that the dynamics are not significantly different in the unchanged and inverted worlds.

the unchanged world) as the experiment progresses.

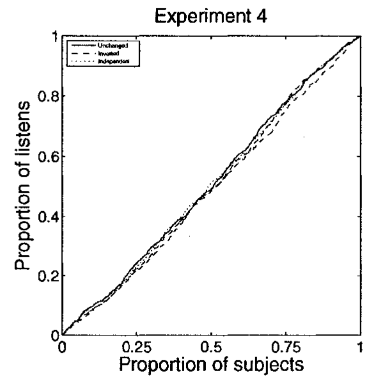
Because we are not fond of data smoothing, we attempted one addition method of looking for patterns over time: comparing the cumulative density plots.[5] Figure D.10(a) plot the cumulative number of listens in each world after a given proportion of subjects. If all subjects listened to the same number of songs, this cumulative plot would fall on the 45-degree line. If earlier subjects listened to more songs the cumulative would be concave, and if later subjects listened to more songs the cumulative would be convex. The striking thing about the the plots, as well as the cumulative download plot (figure D.10(b)) is how closely it follows the 45-degree line in all worlds. Thus, again we fail to have any evidence of subjects in the inverted worlds listening to and downloading more songs over time.

To review, in table 5.4 we found that subjects in the inverted worlds listened to and downloaded few songs. We argued that this was because people in the inverted worlds were disproportionately exposed to lower appeal songs (figure 5.13). However, as time progressed the inverted worlds gradually returned to their original order

[5]This idea is based on the logic of the Kolmogorov-Smirnov test for comparing distributions (Press et al., 1992).



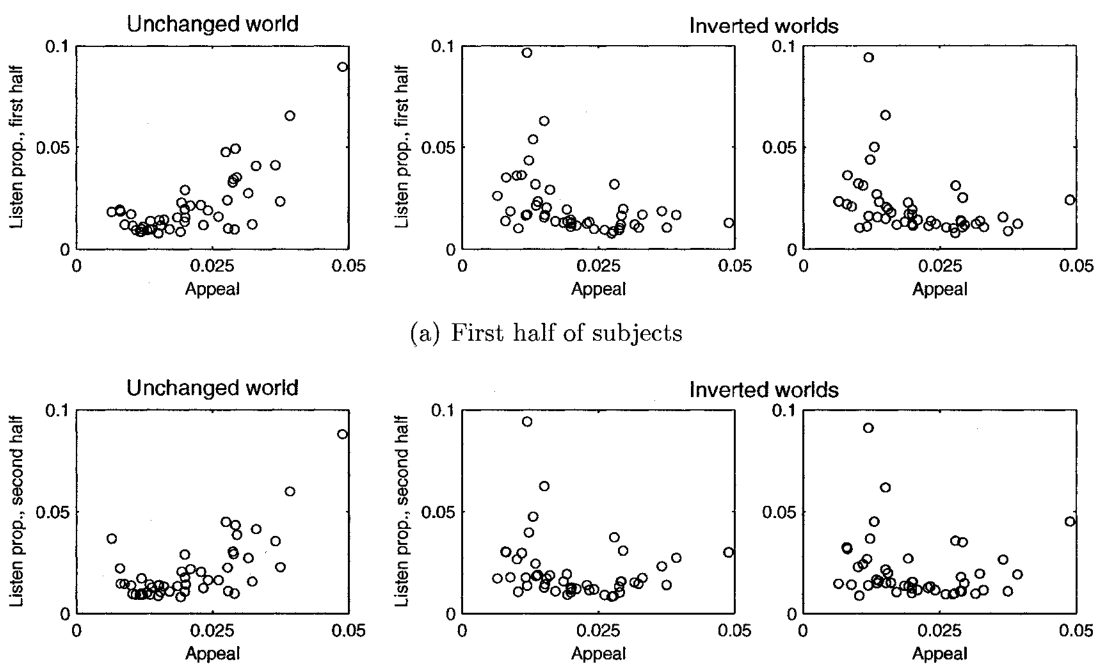
(a) Cumulative listens



(b) Cumulative downloads

Figure D.10: Cumulative number of listens and downloads in experiment 4. If later subjects listened or downloaded at a higher rate, the curve for that world would be concave. Because all curves roughly fall along the 45-degree line, it seems that the dynamics are not significantly different in the different worlds.

(figure 5.21) leading us to speculate that as the worlds returned to their original order, the number of listens and downloads per subject might increase. Despite repeated torturing of the data this does not seem to be the case. One possible reason is that the difference in exposure to bad songs did not change sufficiently. Figure D.11(a) and D.11(b) plot the relationship between appeal and listens in the first and second of the data (the overall relationship is plotted in figure 5.13). In the second half of the experiment the subjects are slightly less overexposed to bad songs, but the difference in minor. Perhaps if the experiment had run longer and the correlation between appeal and success had continue to increase (which might not have happened), then we would have been able to detect increasing number of listens and downloads over time in the inverted worlds. The question of how the total number of downloads (system throughput) can be altered by feedback about the behavior of other subjects may be one that deserves further study because of its obvious practical implications.



(a) First half of subjects

(b) Second half of subjects

Figure D.11: Relationship between appeal and number of listens in first and second half of experiment 4. Subjects in the second half are slightly less overexposed to bad songs, but the difference in not large.

D.4 Notes on timescale of recovery for pairwise comparisons

In the section 5.6 we compared the success of pairs of songs that were swapped by the inversion. We found that in some cases the songs return to the their original ordering, but that in other cases they do not. Here we explore the timescale of recovery; that is, if the songs do return to their original order how many subjects does it take? Figure D.12 plots the results for the top 10 songs before the inversion. Each song has two results (one for each inverted world); circles indicate the songs will return to the their original order while x's indicate that the songs will not. Generally, it seems that the songs that were ranked highest before the inversion were the least likely to recover their original order. Further, if they do return to their original order, it takes a longer time because for these songs the gap to be closed (in terms of downloads) is largest. However, it is the case that the specific value of this timescale is somewhat artificial because it depends on the initial distribution of downloads. For example, if when setting the initial conditions we had multiplied all the download counts by 10 the timescales would all increase by a factor of 10. To conclude, the results seem quite noisy which, as we discussed in the main text, is another reason not to put to much faith in comparing pairs of songs.

D.5 Relationship between the inverted worlds

In previous chapters we have shown that world beginning with the same initial conditions can diverge to different outcomes. For this very reason, we created two inverted worlds, rather than just one. In this experiment, the two inverted worlds are projected to settle into similar steady states (rank correlation = 0.85) (Fig. D.13A). Further, this conclusion is relatively robust to window size (Fig. D.13B). Thus it

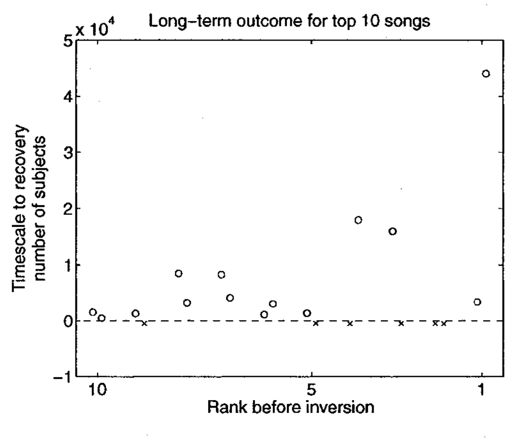


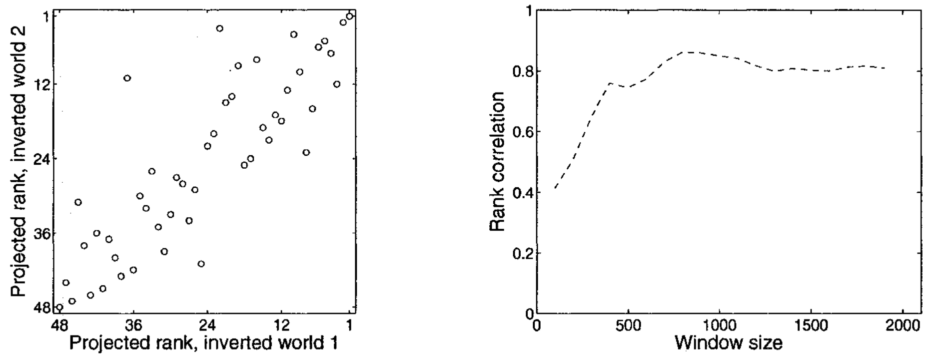
Figure D.12: Timescale to recovery of original ordering for pairwise comparisons.

seems that we can be somewhat confident that we did not observe some extremely unusual come for the inverted worlds.

D.6 A note on the top and bottom 10 songs

One unusual feature of inversion was that there were ties which were broken randomly which in a sense gives a mini-experiment within the larger experiment. For example, there were 10 songs tied for least popular so the top 10 songs all got mapped to the same spot at the bottom (table 5.1). Then it because reasonable to wonder, if, starting from the same initial conditions these songs would again return to the same popularity ordering. Figure D.14 plots the download trajectories of these top 10 songs after the inversion. It turns out that they do not return to the same ordering. For example, in both inverted worlds, song 2 finishes below many of the other songs. Also, in one of the inverted worlds, song 10 earned more downloads than song 1. However, song 1 is rapidly gaining downloads and probably would have eventually overtaken song 10 if the experiment had continued.

Not only where the 10 top songs mapped to the same bottom spot, the 10 bottom songs were mapped to the different spots at the top (table 5.1). To be clear,



(a) Relationship between inverted worlds (b) Robustness of relationship between inverted worlds

Figure D.13: Relationship between inverted worlds. There is a strong relationship between projected final ranks in inverted world 1 and inverted world 2 (rank correlation = 0.85). Further, this correlation is relatively robust to window size.

the label assigned to each of these songs tied at the bottom was determined by the song with which it switched places. For example, song 48 is the one that switched places with song 1, song 47 with song 2, etc. Figure D.15 plots the download trajectories for the 10 initially least popular songs and shows that even thought song 46 and 47 were tied with 9 downloads before the inversion, song 46 overtook song 47 when they were both moved into the top 10. We can also see that song 41 ("Keep your Eyes on the Ballistics" by Secretary) earned almost no downloads (nearly horizontal download trajectory) despite being pushed into the top 10.

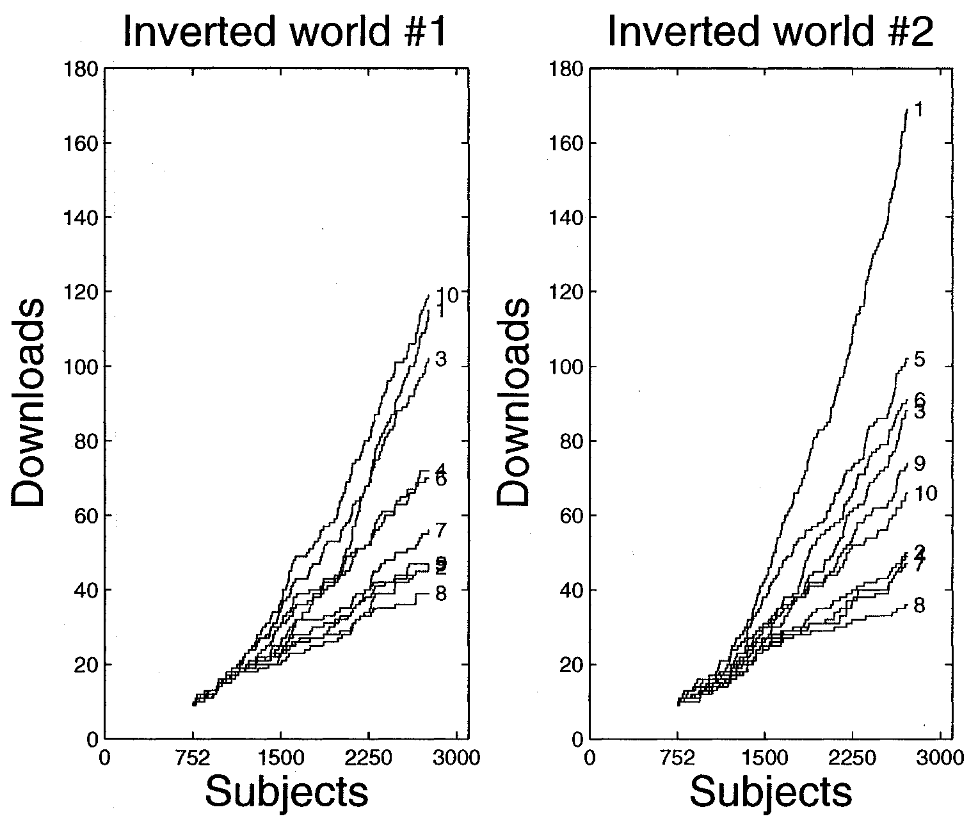


Figure D.14: Success dynamics of the top 10 songs during the set-up period in inverted worlds.

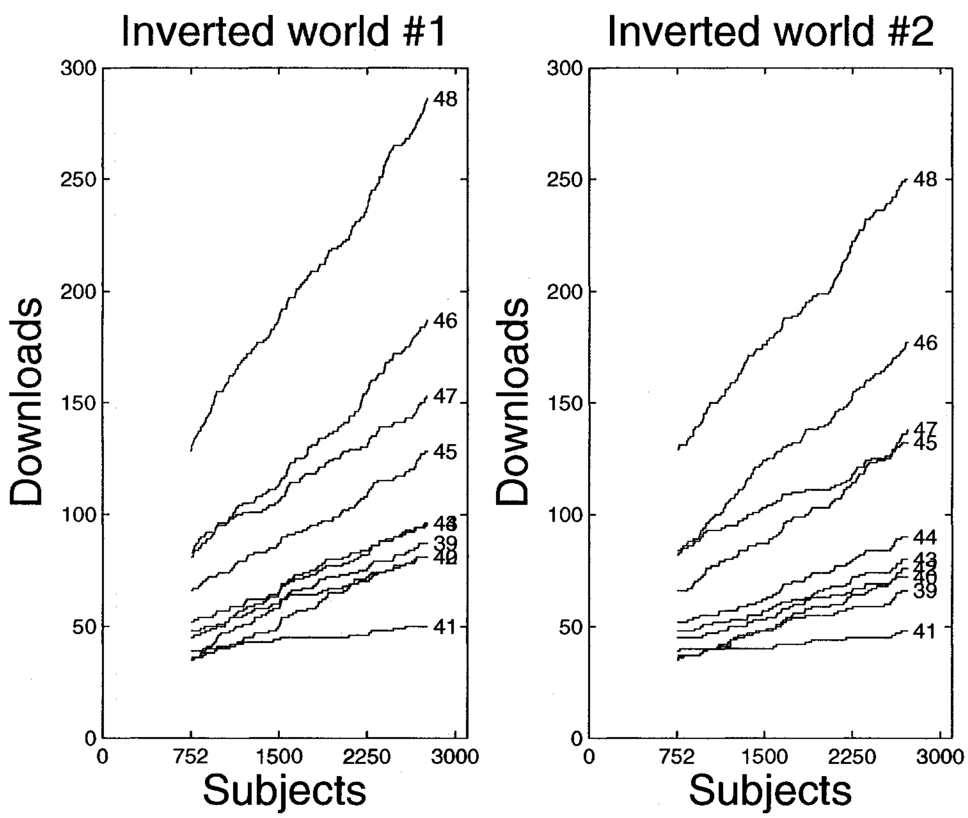


Figure D.15: Success dynamics of the bottom 10 songs during the set-up period in inverted worlds.

Appendix E

Appendix to chapter 6

E.1 Correlation and rank correlation profiles: Additional plots

A number of figures are presented here for archival purposes. These plots show that the results presented in section 6.1 that show that there was lock-in for the top 10 songs are relatively robust. The figures included are:

- Figures E.1, E.2, and E.3 present the correlation profile for the top 5 songs (instead of top 10 songs as was in the main text (figures 6.13, 6.14, and 6.15)).
- Figures E.4, E.5, and E.6 present the rank correlation profile for the top 10 songs (instead of the correlation profile as was in the main text (figures 6.13, 6.14, and 6.15)).
- Figures E.7, E.8, and E.9 present the results where the x-axis is the number of subjects (instead of the proportion of subjects as was in the main text (figures 6.13, 6.14, and 6.15)).

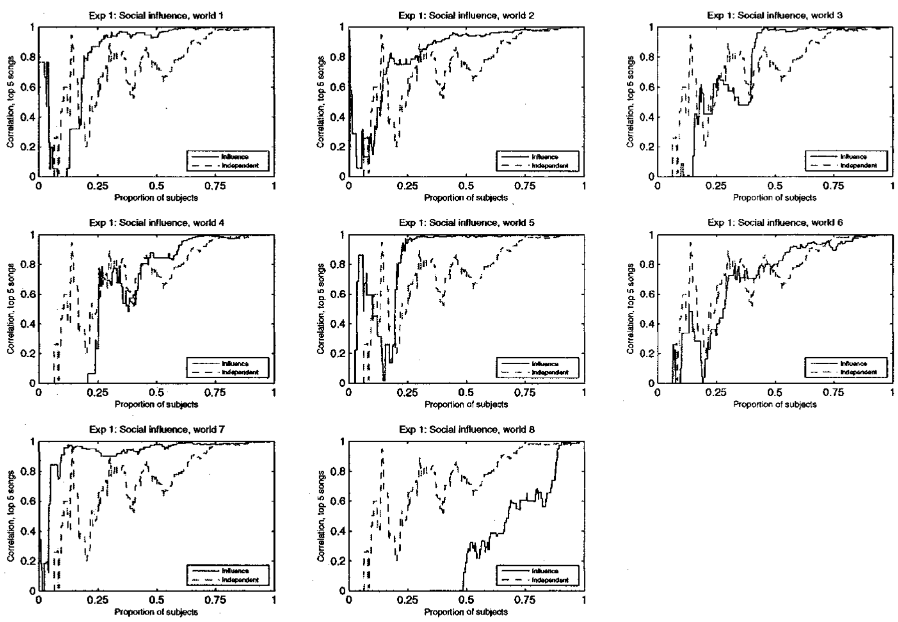


Figure E.1: Correlation profile for the top 5 songs in all social influence worlds in experiment 1 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

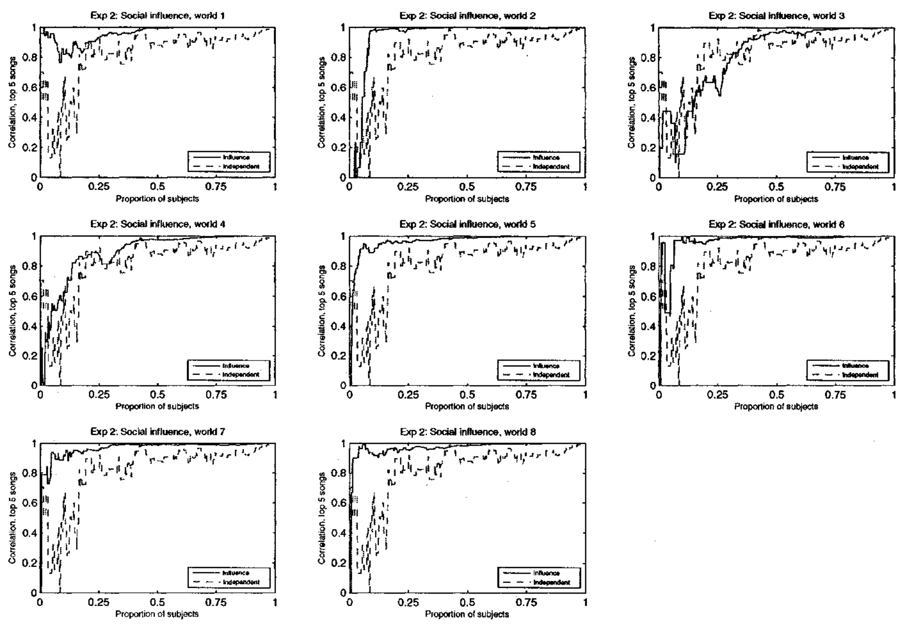


Figure E.2: Correlation profile for the top 5 songs in all social influence worlds in experiment 2 (x-axis = proportion of subjects The dashed line represents the value for the independent condition which can be used as a comparison.

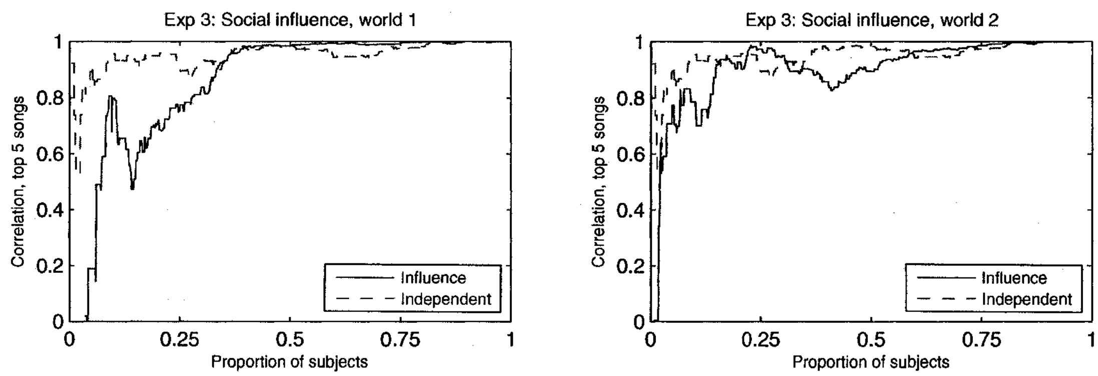


Figure E.3: Correlation profile for the top 5 songs in all social influence worlds in experiment 3 (x-axis = proportion of subjects The dashed line represents the value for the independent condition which can be used as a comparison.

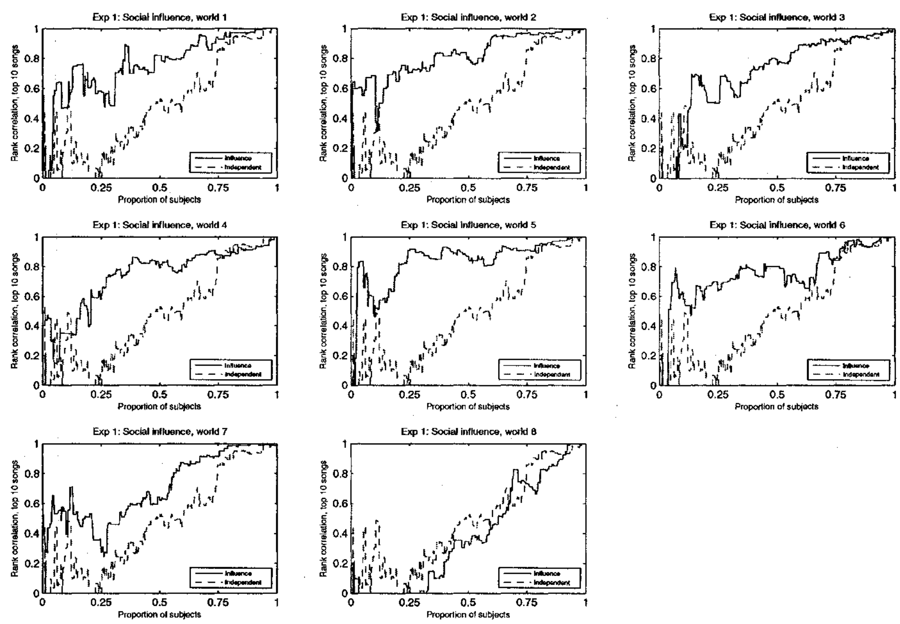


Figure E.4: Rank correlation profile for the top 10 songs in all social influence worlds in experiment 1 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

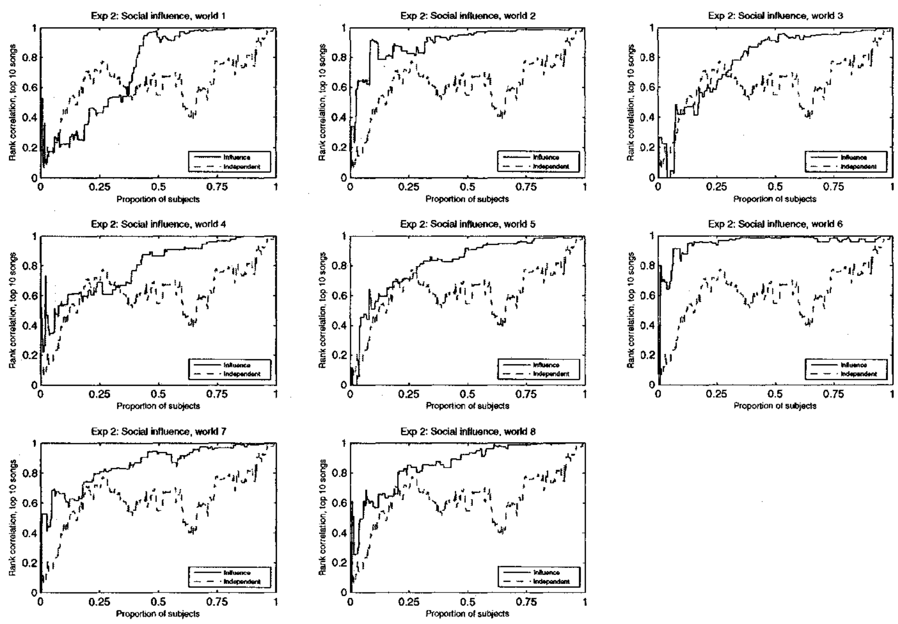


Figure E.5: Rank correlation profile for the top 10 songs in all social influence worlds in experiment 2 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

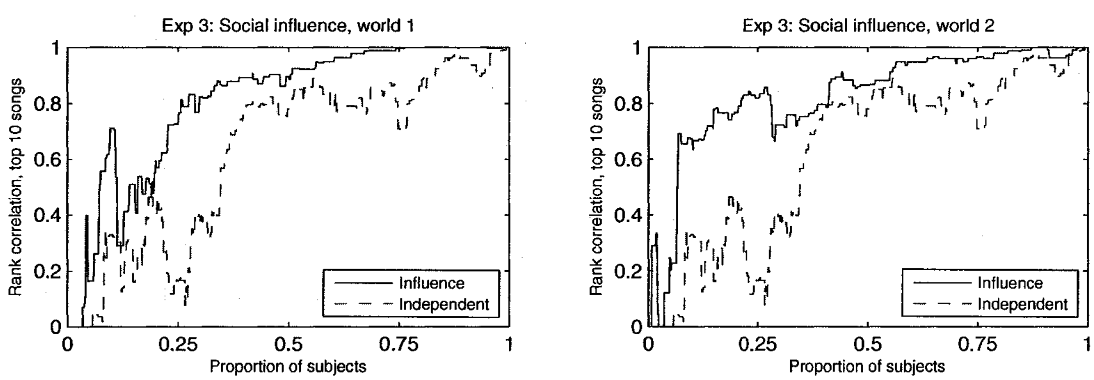


Figure E.6: Rank correlation profile for the top 10 songs in all social influence worlds in experiment 3 (x-axis = proportion of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

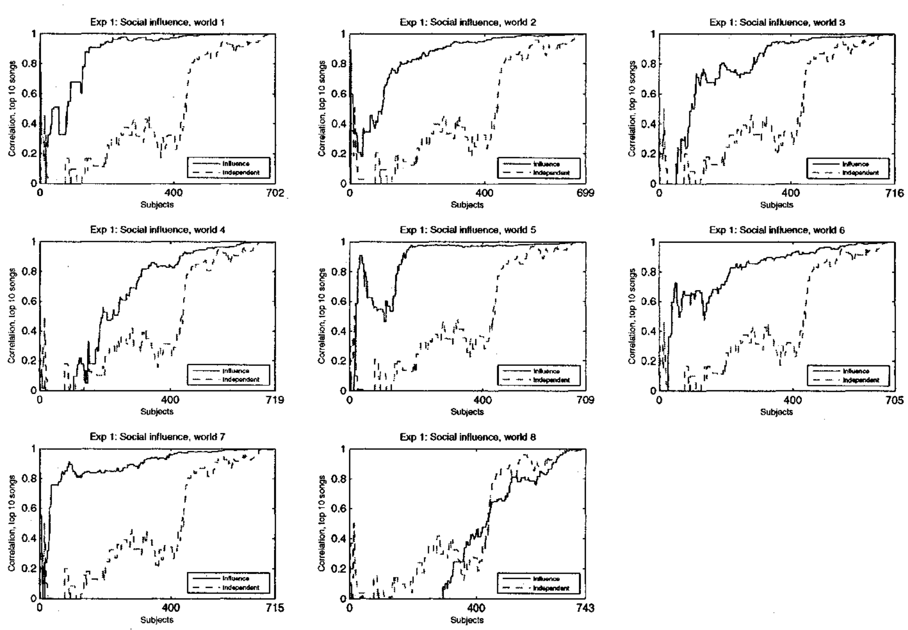


Figure E.7: Correlation profile for the top 10 songs in all social influence worlds in experiment 1 (x-axis = number of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

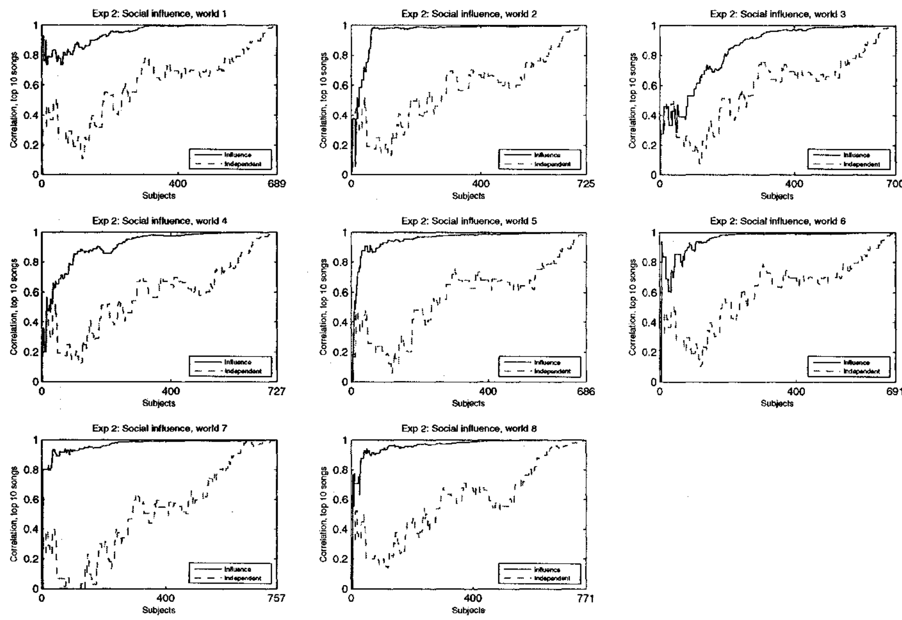


Figure E.8: Correlation profile for the top 10 songs in all social influence worlds in experiment 2 (x-axis = number of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

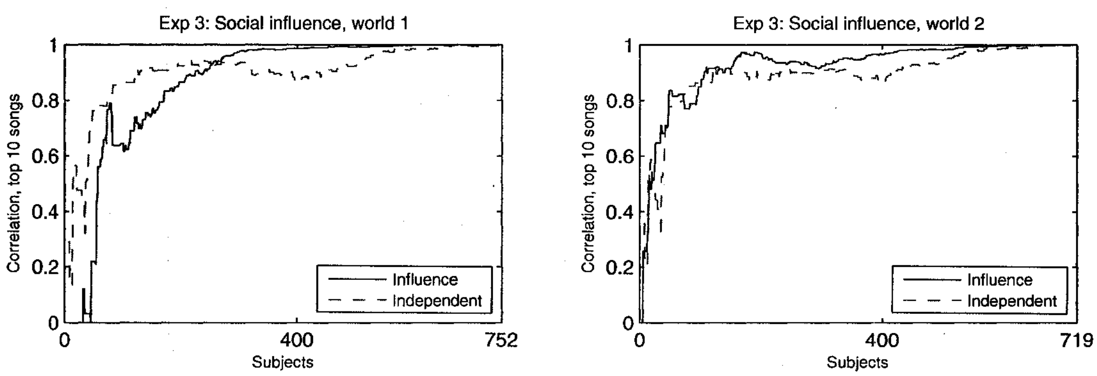
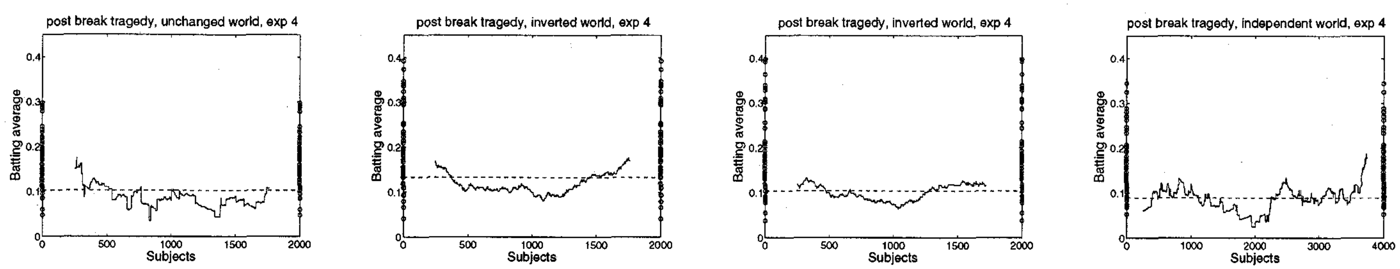


Figure E.9: Correlation profile for the top 10 songs in all social influence worlds in experiment 3 (x-axis = number of subjects). The dashed line represents the value for the independent condition which can be used as a comparison.

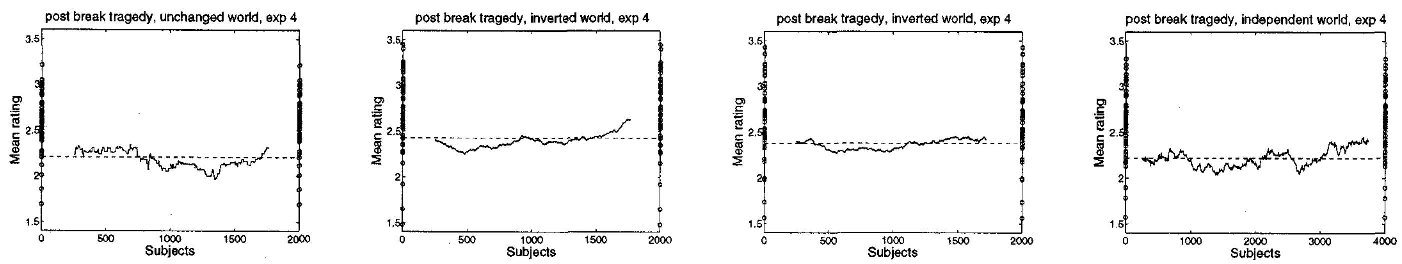
E.2 Was there social influence on the rating and download decision? Additional results

Figures E.10 and E.11 present the dynamics of the batting average and mean rating for song 48 ("Florence" by Post Break Tragedy). Both plots show that the results did not change much during the course of experiment 4.



(a) Unchanged (b) Inverted #1 (c) Inverted #2 (d) Independent

Figure E.10: Dynamics of batting average for (song 48) "Florence" by Post Break Tragedy in the 4 worlds in experiment 4. This song had a higher batting average in the inverted worlds than the unchanged world.



(a) Unchanged (b) Inverted #1 (c) Inverted #2 (d) Independent

Figure E.11: Dynamics of mean rating for "Florence" by Post Break Tragedy in the 4 worlds in experiment 4. This song had a higher mean rating in the inverted worlds than the unchanged world.

Bibliography

Adler, M., 1985. Stardom and talent. American Economic Review 75 (1), 208–212.

Aizen, J., Huttenlocher, D., Kleinberg, J., Novak, A., 2004. Traffic-based feedback on the web. Proceedings of the National Academy of Sciences, USA 101 (Supp 1.), 5254–5260.

Allison, P. D., 1978. Measures of inequality. American Sociological Review 43 (6), 865–880.

Allison, P. D., Long, J. S., Krauze, T. K., 1982. Cumulative advantage and inequality in science. American Sociological Review 47 (5), 615–625.

Alsop, J. A., 1986. The faker's art. New York Review of Books 33 (16).

Anderson, L. R., Holt, C. A., 1997. Information cascades in the laboratory. American Economic Review 87 (5), 847–862.

Arthur, W. B., March 1989. Competing technologies, increasing returns, and lock-in by historical events. The Economic Journal 99 (394), 116–131.

Arthur, W. B., 1994. Increasnig Returns and Path Dependence in the Economy. University of Michigan Press, Ann Arbor.

Arthur, W. B., Ermoliev, Y. M., Kaniovski, Y. M., 1987. Path-dependent processes and the emergence of macro-structure. European Journal of Operational Research 30, 294–303.

Asch, S. E., 1952. Social Psychology. Prentice-Hall, Englewood Cliffs, NJ.

Banerjee, A. V., 1992. A simple model of herd behavior. Quarterly Journal of Economics 107 (3), 797–817.

Barabási, A.-L., Albert, R., 1999. Emergence of scaling in random networks. Science 286 (5439), 509–512.

Bass, F. M., 1969. A new product growth model for consumer durables. Management Science 13 (5), 215–227.

Bass, F. M., Krishnan, T., Jain, D. C., 1994. Why the Bass model fits without decision variables. Marketing Science 13, 203–223.

Baumol, W. J., 1986. Unnatural value: or art investment as a floating crap game. American Economic Review 76 (2), 10–14.

Bearden, W. O., Netemeyer, R. G., Teel, J. E., 1989. Measurement of consumer susceptibiltiy to interpersonal influence. Journal of Consumer Research 15 (4), 473–481.

Bechtel, R., 1967. Hodometer research in museums. Museum News 45 (3), 23–26.

Berger, J., Fisek, M. H., Norman, R. Z., Zelditch, Jr., M., 1977. Status Characteristics and Social Interaction. Elsevier, New York.

Berns, G. S., Chappelow, J., Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Richards, J., 2005. Neurobiological correlates of social conformity and independence during mental rotation. Biological Psychiaty 58 (3), 245–253.

Biel, S., 2005. American Gothic. W.W. Norton, New York.

Bielby, W. T., Bielby, D. D., 1994. 'All hits are flukes': Institutionalized decision making and the rhetoric of network prime-time program development. American Journal of Sociology 99 (5), 1287–1313.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashions, customs, and cultural change as information cascades. Journal of Political Economy 100 (5), 992–1026.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1998. Learning from the behavior of others: Conformity, fads, and information cascades. Journal of Economic Perspectives 12 (3), 151–170.

Birnbaum, M. H., 2004. Human research and data collection via the internet. Annual Review of Psychology 55, 803–832.

Bonabeau, E., 2004. The perils of the imitation age. Harvard Business Review 82 (6), 45–54.

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

Bornstein, B. H., 1999. The ecological validity of jury simulations: Is the jury still out? Law and Human Behavior 23 (1), 75–91.

Bornstein, R. F., 1989. Exposure and affect: Overview and meta-analysis of research, 1968-1987. Psychological Bulletin 106 (2), 265–289.

Bourdieu, P., 1984. Distinction: A Social Critique of the Judgment of Taste. Harvard University Press, Cambridge, MA.

Bradley, R. A., 1976. Science, statistics, and paired comparisons. Biometrics 32 (2), 213–232.

Bruch, E. E., Mare, R. D., 2006. Neighborhood choice and neighborhood change. American Journal of Sociology 112 (3), 667–709.

Brynjolfsson, E., Hu, Y., Smith, M. D., 2003. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. Management Science 49 (11), 1580–1596.

Burleson, B. R., Levine, B. J., Samter, W., 1984. Decision-making procedure and decision quality. Human Communication Research 10 (4), 557–574.

Carter, B., 2006. Desperate Networks. Doubleday, New York.

Caves, R. E., 2000. Creative Industries: Contracts between art and commerce. Harvard University Press, Cambridge.

Chase, I. D., Tovey, C., Sprangler-Martin, D., Manfredonia, M., 2002. Individual differences versus social dynamics in the formation of animal dominance hierarchies. Proceedings of the National Academy of Sciences, USA 99 (8), 5744–5749.

Chevalier, J. A., Mayzlin, D., 2006. The effect of work of mouth on sales: Online book reviews. Journal of Marketing Research XLIII, 345–354.

Chung, K. H., Cox, R. A. K., 1994. A stochastic model of superstardom: An application of the Yule distribution. The Review of Economics and Statistics 76 (4), 771–775.

Cialdini, R. B., 2001. Influence: Science and Practice, 4th Edition. Allyn and Bacon, Boston.

Cialdini, R. B., Goldstein, N. J., 2004. Social influence: Compliance and conformity. Annual Review of Psychology 55, 591–621.

Coase, R. H., 1979. Payola in radio and television broadcasting. Journal of Law and Economics 22 (2), 269–328.

Coleman, J. S., 1990. Foundations of Social Theory. Harvard University Press, Cambridge, MA.

Colley, A., North, A., Hargreaves, D. J., 2003. Gender bias in the evaluation of New Age music. Scandinavian Journal of Psychology 44, 125–131.

Considine, S., 2004. The hit we almost missed. The New York Times (Dec. 3).

Cook, K. S., Emerson, R. M., Gillmore, M. R., Yamagishi, T., 1983. The distribution of power in exchange networks - theory and experimental results. American Journal of Sociology 89 (2), 275–305.

Costner, H. L., 1965. Criteria for measures of association (with reply). American Sociological Review 30 (3), 341–353.

Coulter, P. B., 1989. Measuring Inequality: A Methodological Handbook. Westview Press, Boulder.

Cutting, J. E., 2003. Gustave Caillebotte, French Impressionism, and mere exposure. Psychonomic Bulletin & Review 10 (2), 319–343.

David, H. A., 1988. The Method of Paired Comparisons. Oxford University Press, New York.

David, P., 1985. Clio and the economics of QWERTY. American Economic Review 75 (2), 332–337.

Dawes, R., 1980. Social dilemmas. Annual Review of Psychology 31, 169–193.

de Solla Price, D., 1976. A general theory of bibliometric and other cumulative advantage processes. Journal of the American Society for Information Science 27 (5), 292–306.

De Vany, A., 2004. Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry. Routledge, New York.

De Vany, A., Lee, C., 2001. Quality signals in information cascades and the dynamics of the distribution of motion picture box office revenues. Journal of Economic Dynamics and Control 25, 593–614.

De Vany, A., Walls, W. D., 1996. Bose-Einstein dynamics and adaptive contracting the motion picture industry. The Economic Journal 106 (439), 1493–1514.

De Vany, A., Walls, W. D., 1997. The market for motion pictures: Rank, revenue and survival. Economic Inquiry 35 (4), 783–797.

De Vany, A., Walls, W. D., 1999. Uncertainty in the movie industry: Does star power reduce the terror of the box office? Journal of Cultural Economics 23, 285–318.

De Vany, A., Walls, W. D., 2004. Motion picture profit, the stable Paretian hypothesis, and the curse of the superstar. Journal of Economic Dynamics and Control 28, 1035–1057.

de Vries, H., 1998. Finding a dominance order most consistent with a linear hierarchy: Procedure and review. Animal Behavior 55, 827–843.

Dellarocas, C., October 2006. Strategic manipulation of Internet opinion forums: Implications for consumers and firms. Management Science 52 (10), 1577–1593.

Denisoff, R. S., 1975. Solid Gold: The Popular Record Industry. Transaction Books, New Brunswick.

Deutsch, M., Gerard, H. B., 1955. A study of normative and informative social influences upon individual judgement. Journal of Abnormal Social Psychology 51, 629–636.

Diamond, S. S., 1997. Illuminations and shadows from jury simulations. Law and Human Behavior 21 (5), 561–571.

DiMaggio, P., 1987. Classification in art. American Sociological Review 52 (4), 440–455.

DiPrete, T. A., Eirich, G., 2006. Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. Annual Review of Sociology 32, 271–297.

Dittrich, R., Hatzinger, R., Katzenbeisser, W., 1998. Modelling the effect of subject-specific covariates in paired comparison studies with an application to university rankings. Applied Statistics 47 (4), 511–525.

Dodd, S. C., 1956. A test of message diffusion by chain tags. American Journal of Sociology 61 (5), 425–432.

Dodds, P. S., Muhamad, R., Watts, D. J., 2003. An experimental study of search in global social networks. Science 301, 827–829.

Dodds, P. S., Watts, D. J., 2004. Universal behavior in a generalized model of contagion. Physical Review Letters 92, 218701.

Dodds, P. S., Watts, D. J., 2005. A generalized model of social and behavioral contagion. Journal of Theoretical Biology 232, 587–604.

Dowd, T. J., Liddle, K., Lupo, K., Borden, A., 2002. Organizing the musical canon: The repertoires of major U.S. symphony orchestras, 1842 to 1969. Poetics 30, 35–61.

Elman, C., O'Rand, A. M., 2004. The race is to the swift: Socioeconomic origins, adult education, and wage attainment. American Journal of Sociology 110 (1), 123–160.

Faulkner, R. R., Anderson, A. B., 1987. Short-term projects and emergent careers: Evidence from Hollywood. American Journal of Sociology 92 (4), 879–909.

Felzenszwalb, P. F., Huttenlocher, D. P., Kleinberg, J. M., 2004. Fast algorithms for large-state-space hmms with applications to web usage analysis. In: Thrun, S., Saul, L., Schölkopf, B. (Eds.), Advances in Neural Information Processing Systems 16. MIT Press, Cambridge, MA.

Fienberg, S. E., Larntz, K., 1976. Log linear representation for paired and multiple comparisons models. Biometrika 63 (2), 245–254.

Fisher, M., 2007. Something in the Air: Radio, Rock, and the Revolution that Shaped a Generation. Random House, New York.

Fox, M. A., Kochanowski, P., 2004. Models of superstardom: An application of the Lotka and Yule distributions. Popular Music and Society 27 (4), 507–522.

Frank, R. H., Cook, P. J., 1995. The Winner-Take-All Society. The Free Press, New York.

Fuller, R. G. C., Sheehy-Skeffington, A., 1974. Effect of group laughter on responses to humorous materials: A replication and extension. Psychological Reports 35, 531–534.

Gans, H. J., 1974. Popular Culture and High Culture: An Analysis and Evaluation of Taste. Basic Books, New York.

Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B., 2004. Bayesian Data Analysis, 2nd Edition. Chapman & Hall, Boca Raton.

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York.

Gilmore, S., 1993. Tradition and novelty in concert programming: Bringing the artist back into cultural analysis. Sociological Forum 8 (2), 221–242.

Ginsburgh, V., Spring 2003. Awards, success, and aesthetic quality in the arts. Journal of Economic Perspectives 17 (2), 99–111.

Gitlin, T., 1983. Inside Prime Time. Pantheon Books, New York.

Gladwell, M., October 16 2006. The formula. The New Yorker , 138–149.

Godes, D., Mayzlin, D., Chen, Y., Das, S., Dellarocas, C., Pfeiffer, B., Libai, B., Sen, S., Shi, M., Verlegh, P., 2005. The firm's management of social interactions. Marketing Letters 16, 415–428.

Goldman, W., 1983. Adventures in the Screen Trade. Warner Books, New York.

Goldstein, D. G., Gigerenzer, G., 2002. Models of ecological rationality: The recognition heuristic. Psychological Review 109 (1), 75–90.

Goodman, L. A., Kruskal, W. H., 1954. Measures of association for cross classifications. Journal of the American Statistical Association 49 (268), 732–764.

Gordon, M. E., Slade, L. A., Schmitt, N., 1986. The "science of the Sophomore" revisted: From conjecture to empiricism. Academy of Management Review 11 (1), 191–207.

Gosling, S. D., Vazire, S., Srivastava, S., John, O. P., 2004. Should we trust web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. American Psychologist 59 (2), 93–104.

Gould, R. V., 2002. The origins of status hierarchies: A formal theory and empirical test. American Journal of Sociology 107 (5), 1143–1178.

Greco, A., 1997. The Book Publishing Industry. Allyn & Bacon.

Hamlen, Jr., W. A., 1991. Superstardom in popular music: Empirical evidence. The Review of Economics and Statistics 73 (4), 729–733.

Hamlen, Jr., W. A., 1994. Variety and superstardom in popular music. Economics Inquiry 32 (3), 395–406.

Hanson, W. A., Putler, D. S., 1996. Hits and misses: Herd behavior and online product popularity. Marketing Letters 7 (4), 2970395.

Heath, C., Ho, B., Berger, J., 2006. Focal points in coordinated divergence. Journal of Economic Psychology 27, 635–647.

Heckathorn, D. D., 1997. Respondent-driven sampling: A new approach to the study of hidden populations. Social Problems 44 (2), 174–199.

Heckathorn, D. D., 2002. Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. Social Problems 49 (1), 11–34.

Heckman, J. J., 2006. Skill formation and the economics of investing in disadvantaged children. Science 312, 1900–1902.

Hedström, P., 1998. Rational imitation. In: Hedström, P., Swedberg, R. (Eds.), Social Mechanisms: Analytical Approach to Social Theory. Cambridge University Press, Cambridge, pp. 306–327.

Hedström, P., 2005. Dissecting the Social: On the Principles of Analytical Sociology. Cambridge University Press, Cambridge.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F. W., Patton, J. Q., Tracer, D., 2005. "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. Behavioral and Brain Science 28, 795–855.

Hill, B., Power, D., 2005. The Making of a Bestseller: Success Stories from Authors and the Editors, Agents, and Booksellers Behind Them. Dearborn, Chicago.

Hirsch, P. M., 1972. Processing fads and fashions: An organization-set analysis of cultural industry systems. American Journal of Sociology 77 (4), 639–659.

Hirschberg, L., 2005. Giving them what they want. The New York Times Magazine (Sep. 4).

Holland, P. W., 1986. Statistics and causal inference. Journal of the American Statistical Association 81 (396), 945–960.

Horowitz, I. A., Bordens, K. S., 2002. The effects of jury size, evidence complexity, and note taking on jury process and performance in a civil trial. Journal of Applied Psychology 87 (1), 121–130.

Hoyer, W. D., Brown, S. P., 1990. Effects of brand awareness on choice for a common, repeat purchase product. Journal of Consumer Research 17 (2), 141–148.

Huang, J.-H., Chen, Y.-F., 2006. Herding in online product choice. Psychology & Marketing 23 (5), 413–428.

Jussim, L., 1986. Self-fulfilling prophecies: A theoretical and integrative review. Psychological Review 93 (4), 429–445.

Kaplan, M. F., Miller, C. E., 1987. Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. Journal of Personality and Social Psychology 53 (2), 306–313.

Katz, E., Lazarsfeld, P. F., 1955. Personal Influence: The part played by people in the flow of mass communications. The Free Press, Glencoe.

Kendall, M., Gibbons, J. D., 1990. Rank Correlation Methods, 5th Edition. Edward Arnold, London.

Klein, H.-J., 1993. Tracking visitor circulation in museum settings. Environment and Behavior 25 (6), 782–800.

Krapivsky, P., Redner, S., 2001. Organization of growing random network. Physical Review E 63, 066123.

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., Couper, M., 2004. Psychological research online: Report of board of scientific affairs' advisory group on the conduct of research on the Internet. American Psychologist 59 (2), 105–117.

Krueger, A. B., 2005. The economics of real superstars: The market for rock concerts in the material world. Journal of Labor Economics 23 (1), 1–30.

Kruskal, W. H., 1958. Ordinal measures of association. Journal of the American Statistical Association 53 (284), 814–861.

Kviz, F. J., 1981. Interpreting proportional reduction in error measures as percentage of variation explained. The Sociological Quarterly 22, 413–420.

Laureti, P., Moret, L., Zhang, Y.-C., 2005. Aggregating partial, local evaluations to achieve global rankings. Physica A 345, 705–712.

Lawler, E. J., Yoon, J., 1996. Commitment in exchange relations: Test of a theory of relational cohesion. American Sociological Review 61 (1), 89–108.

Lawless, J., 2005. The interview: Nigel Newton: Is there life after Harry? You bet your Hogwarts there is. Independent (London) (Jul. 3).

Lee, S. E., 1986. The faker's art. New York Review of Books 33 (20).

Liebowitz, S. J., Margolis, S. E., 1995. Path dependence, lock-in, and history. Journal of Law, Economics, & Organization 11 (1), 205–226.

Liedtke, W., Logan, C., Orenstein, N. M., Dickey, S. S., 1995. Rembrandt/Not Rembrandt in The Metropolitan Museum of Art: Aspects of Connoisseurship. Vol. II: Paintings, Drawings, and Prints: Art-Historical Perspectives. Metropolitan Museum of Art, New York.

Lovaglia, M. J., Lucas, J. W., Houser, J. A., Thye, S. R., Markovsky, B., 1998. Status processes and mental ability test scores. American Journal of Sociology 104 (1), 195–228.

Lucas, J. W., 2003. Theory-testing, generalization, and the problem of external validity. Sociological Theory 21 (3), 236–253.

MacDonald, G. M., 1988. The economics of rising stars. American Economic Review 78 (1), 155–166.

Mahoney, J., 2000. Path dependence in historical sociology. Theory and Society 29, 507–548.

Merton, R. K., 1936. The unanticipated consequences of purposive social action. American Sociological Review 1 (6), 894–904.

Merton, R. K., 1948. The self-fulfilling prophecy. Antioch Review 8 (2), 193–210.

Merton, R. K., 1968. The Matthew effect in science. Science 159 (3810), 56–63.

Merton, R. K., 1995. The Thomas theorem and the Matthew effect. Social Forces 74 (2), 379–422.

Milgram, S., Bickman, L., Berkowitz, L., 1969. Note on the drawing power of crowds of different size. Journal of Personality and Social Psyhcology 13 (2), 79–82.

Mintz, A., Redd, S. B., Vedlitz, A., 2006. Can we generalize from student experiments to the real world in political science, military affairs, and international relations? Journal of Conflict Resolution 50 (5), 757–776.

Molm, L. D., Takahashi, N., Peterson, G., 2000. Risk and trust in social exchange: An experimental test of a classical proposition. American Journal of Sociology 105 (5), 1396–1427.

Newman, M. E. J., 2003. The structure and function of complex networks. SIAM Review 45 (2), 167–256.

North, A. C., Colley, A. M., Hargreaves, D. J., 2003. Adolescents' perceptions of the music of male and female composers. Psychology of Music 31 (2), 139–154.

North, A. C., Hargreaves, D. J., 1997. Experimental aesthetics and everyday music listening. In: Hargreaves, D. J., North, A. C. (Eds.), The Social Psychology of Music. Oxford University Press, Oxford, Ch. 5, pp. 84–103.

Nosek, B. A., Banaji, M. R., Greenwald, A. G., 2002. E-research: Ethics, security, design, and control in Psyhchological research on the Internet. Journal of Social Issues 58 (1), 161–176.

Ohmer, S., 2006. George Gallup in Hollywood. Columbia University Press, New York.

Ordonez, J., 2002. Behind the music: MCA spent millions on Carly Hennessy – haven't heard of her? The Wall Street Journal (Feb. 26).

Page, S. E., 2006. Path dependence. Quarterly Journal of Political Science 1, 87–115.

Parish, J. R., 2006. Fiasco: A History of Hollywood's Iconic Flops. John Wiley & Sons, Hoboken, NJ.

Park, C. W., Lessig, V. P., 1977. Students and housewives: Difference in susceptibility to reference group influence. Journal of Consumer Research 4 (2), 102–110.

Pasupathi, M., 1999. Age differences in response to conformity pressure for emotional and nonemotional material. Psychology and Aging 14 (1), 170–174.

Peterson, K., 1997. The distribution and dynamics of uncertainty in art galleries: A case study of new dealerships in the Parisian art market, 1985-1990. Poetics 25 (4), 241–263.

Peterson, R. A., 2001. On the use of college students in social science research: Insights from a second-order meta-analysis. Journal of Consumer Research 28, 450–461.

Peterson, R. A., Anand, N., 2004. The production of culture perspective. Annual Review of Sociology 30, 311–334.

Peterson, R. A., Berger, D. G., 1971. Entrepreneurship in organizations: Evidence from the popular music industry. Administrative Science Quarterly 16 (1), 97–106.

Pierson, P., 2000. Increasing returns, path dependence, and the study of politics. American Political Science Review 94 (2), 251–267.

Portes, A., 2000. The hidden abode: Sociology as analysis of the unexpected. American Sociological Review 65 (1), 1–18.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., 1992. Numerical Recipes in C: The Art of Scientific Computing, 2nd Edition. Cambridge University Press.

Randle, W., 1961a. Foreign usage of 'payola'. American Speech 36 (4), 275–277.

Randle, W., 1961b. Payola. American Speech 36 (2), 104–116.

Ridgeway, C. L., 1982. Status in groups: The importance of motivation. American Sociological Review 47 (1), 76–88.

Robinson, M. D., Montgomery, S. S., 2000. The time allocation and earnings of artists. Industrial Relations 39 (3), 525–534.

Rosen, S., 1981. The economics of superstars. American Economic Review 71 (5), 845–858.

Russell, P. A., 1986. Experimental aesthetics of popular music recordings: Pleasingness, familiarity, and chart performance. Psychology of Music 14, 33–43.

Russell, P. A., 1987. Effects of repetition on the familiartity and liekability of popular music recordings. Psychology of Music 15, 187–197.

Salganik, M. J., Dodds, P. S., Watts, D. J., 2006. Experimental study of inequality and unpredictability in an artificial cultural market. Science 311, 854–856.

Salganik, M. J., Heckathorn, D. D., 2004. Sampling and estimation in hidden populations using respondent-driven sampling. Sociological Methodology 34, 193–239.

Sassoon, D., 2001. Becoming Mona Lisa: The making of a global icon. Harcourt, New York.

Sassoon, D., 2006. Leonardo and the Mona Lisa Story: The History of a Painting Told in Pictures. Overlook Duckworth, New York.

Schelling, T. C., 1978. Micromotives and Macrobehaviors. W.W. Norton, New York.

Schkade, D., Sunstein, C. R., Kahneman, D., 2000. Deliberating about dollars: The severity shift. Columbia Law Review 100 (4), 1139–1175.

Schweinhart, L. J., Mortie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., Mores, M., 2005. Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40. High/Scope Educational Research Foundation, Ypsilanti, MI.

Seabrook, J., 2000. Nobrow: The culture of marketing and the marketing of culture. Vintage Books, New York.

Sears, D. O., 1986. College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. Journal of Personality and Social Psychology 51 (3), 515–530.

Senecal, S., Nantel, J., 2004. The influence of online product recommendations on consumers' online choices. Journal of Retailing 80, 159–169.

Sered, S. S., Fernandopulle, R., 2005. Uninsured in America: Life & Death in the Land of Opportunity. University of California Press, Berkeley.

Sherif, M., 1937. An experimental approach to the study of attitudes. Sociometry 1 (1/2), 90–98.

Silverstein, D. A., Farrell, J. E., 2001. Efficent method for paired comparison. Journal of Electronic Imaging 10 (2), 394–398.

Simmel, G., 1957. Fashion. American Journal of Sociology 62 (6), 541–558.

Simon, H. A., 1955. On a class of skew distribution functions. Biometrika 42 (3/4), 425–440.

Simpson, B. S., Macy, M. W., 2004. Power, identity, and collective action in social exchange. Social Forces 82 (4), 1373–1409.

Skitka, L., Sargis, E., 2005. Social psychological research and the Internet: the promise and peril of a new methodlogical frontier. In: Amichai-Hamburger, Y. (Ed.), The Social Net: Human Behavior in Cyberspace. Oxford University Press, Oxford, Ch. 1, pp. 1–25.

Snow, D. A., Anderson, L., 1993. Down on Their Luck: A Study of Homeless Street People. University of California Press, Berkeley.

Sobel, M. E., 1996. An introduction to causal inference. Sociological Methods and Research 24, 353–379.

Somers, R. H., 1962. A new asymmetric measure of association for ordinal variables. American Sociological Review 27 (6), 799–811.

Sorensen, A., 2007. Bestseller lists and product variety. Journal of Industrial Economics , (forthcoming).

Stern, W., August 7 1995. Did dirty tricks create a best-seller? Business Week , 22.

Strobl, E. A., Tucker, C., 2000. The dynamics of chart success in the U.K. pre-recorded popular music industry. Journal of Cultural Economics 24, 113–134.

Trachtenberg, J. A., 2006. In era of blockbuster books one publisher rolls the dice. The Wall Street Journal (Oct. 16).

UNDP, 2004. Human Development Report 2004. UNDP (United Nations Development Programme), New York.

Velthuis, O., 2005. Talking Prices: Symbolic Meanings of Prices on the Market for Contemporary Art. Princeton University Press, Princeton.

Vogel, H. L., 2004. Entertainment Industry Economics: A Guide for Financial Analysis, 6th Edition. Cambridge University Press, Cambridge.

von Sonnenburg, H., 1995. Rembrandt/Not Rembrandt in The Metropolitan Museum of Art: Aspects of Connoisseurship. Vol. I: Paintings: Problems and Issues. Metropolitan Museum of Art, New York.

Wapshott, N., 2005. Harry Potter and the half-based print run. Independent (London) (Oct. 30).

Watts, D. J., 2002. A simple model of global cascades in random networks. Proceedings of the National Academy of Sciences, USA 99 (9), 5766–5771.

Watts, D. J., 2003. Six Degrees: The Science of a Connected Age. W.W. Norton, New York.

Watts, D. J., Peretti, J., May 2007. Viral marketing for the real world. Harvard Business Review , 22–23.

White, H., White, C., 1965. Canvasses and Careers: Institutional Change in the French Painting World. John Wiley, New York.

Willer, D. (Ed.), 1999. Network Exchange Theory. Praeger.

Willer, R., 2005. The intelligibility of unintelligible texts: The effects of author's status. Working Paper .

Williams, A., 2005. Feeding the hunger for the next diet book. The New York Times (Apr. 17).

Winship, C., Morgan, S. L., 1999. The estimation of causal effects from observational data. Annual Review of Sociology 25, 659–706.

Yule, G. U., 1925. A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis, F.R.S. Philosophical Transactions of the Royal Society of London, Series B 213, 21–87.

Zelditch, Jr., M., 1969. Can you really study an army in the laboratory? In: Etzioni, A. (Ed.), A Sociological Reader on Complex Organization, 2nd Edition. Holt, Rinehart, and Winston, New York, pp. 528–539.