Available online at www.sciencedirect.com

ScienceDirect

Journal of Interactive Marketing 28 (2014) 16 – 25

JOURNAL OF INTERACTIVE MARKETING

www.elsevier.com/locate/intmar

Patterns of Herding and their Occurrence in an Online Setting

David J. Langley [a,*] & Maarten C. Hoeve [b] & J. Roland Ortt [c] & Nico Pals [a] & Bob van der Vecht [d]

[a] TNO Netherlands Organisation for Applied Scientific Research, Groningen, Netherlands
[b] TNO Netherlands Organisation for Applied Scientific Research, Delft, Netherlands
[c] Delft University of Technology, Delft, Netherlands
[d] TNO Netherlands Organisation for Applied Scientific Research, The Hague, Netherlands

Available online 26 September 2013

Abstract

When groups of consumers share information or express their opinions about products and services, their attitudes or behavior sometime align without centralized coordination, a phenomenon known as herding. Building on pattern-based explanations of herding from the cognitive science literature, we propose a framework to elucidate herding behavior based on three dimensions: the *speed of contagion*, i.e., the extent to which the behavior spreads in a given time, the *number of individuals*, i.e., the proportion of the whole population expressing the behavior, and the *uniformity of direction*, i.e., the extent to which the mass behavior is increasingly uniform with one variant becoming dominant. Based on these dimensions, we differentiate eight patterns of herding behavior from slowly diffusing, small and disparate groups through to rapidly spreading, massive herds expressing a convergent behavior. We explore these herding patterns in an online setting, measuring their prevalence using over four thousand streams of data from the online micro-blogging application, Twitter. We find that all eight patterns occur in the empirical data set although some patterns are rare, particularly those where a convergent behavior rapidly spreads through the population. Importantly, those occurrences that develop into the pattern we call "stampeding," i.e., the rapid spread of a dominant opinion expressed by many people, generally follow a consistent development path. The proposed framework can help managers to identify such noteworthy herds in real time, and represents a first step in anticipating this form of group behavior.

Keywords: Consumer behavior; Social contagion; Diffusion; Social influence; Social media; Twitter

Introduction

This paper explores dynamic patterns of mass consumer behavior, leading to the phenomenon known as "herding." In this paper, we will focus in particular on herding behavior via online media.

A better understanding of online herding is needed if we are to understand the link between widespread online opinion formation and marketing outcomes, such as sales or reactions to promotion campaigns. Obviously, not all herding in the online setting is positive for firms, with a classic example being the bank run (Gu 2011). In 2008, there was a run on the Icelandic bank, Landsbanki, which had offered high interest online savings accounts in the UK and The Netherlands under the brand, IceSave. Rumors quickly circulated online that the bank had run into financial difficulty and the speed with which online withdrawals were made prompted the Icelandic government to step in to save the nation's banking system.

This paper has two objectives: first, to design a framework to distinguish between different patterns of herding behavior; second, to assess which of these patterns can be witnessed in practice using a large database of online behavior, and to investigate the dynamics of the development of large herds. For example, does a particular sequence of patterns indicate an impending rush of consumer sentiment or behavior? In short, we propose a framework to distinguish between different patterns of mass behavior and identify these patterns using online data.

There is as yet no accepted way of scientifically assessing whether a change in the pattern of online mass behavior is important. One change may prove to be insignificant whereas another may lead to a sudden explosion of consumer reactions.

* Corresponding author.
E-mail addresses: david.langley@tno.nl (D.J. Langley), maarten.hoeve@tno.nl (M.C. Hoeve), j.r.ortt@tudelft.nl (J.R. Ortt), nico.pals@tno.nl (N. Pals), bob.vandervecht@tno.nl (B. van der Vecht).

1094-9968/$ -see front matter
http://dx.doi.org/10.1016/j.intmar.2013.06.005

Hype cycles can be assessed by tracking the number of news items regarding a specific topic, for example, but the resulting pattern with peaks in news items over time is usually erratic and hard to diagnose (Jun 2012). Despite the availability of real-time, online data, there is a lack of well-grounded instruments for modeling changes in mass consumer attitude and behavior.

Besides this scientific motivation, this study is also clearly relevant to managers. Current marketing practice is to track certain characteristics of mass consumer behavior, without combining individual indicators in a coherent framework as a basis for efficient intervention. Firms are able to measure the number of page views of a new campaign, the prevailing sentiment in an online discussion, and they can see when a relevant topic is trending on Twitter. However, it is important to know which metrics to follow simultaneously, and to be able to identify different patterns of mass behavior so that irrelevant noise can be confidently ignored and time and energy can be devoted to situations which signify noteworthy developments. Consumer complaints can occasionally lead to considerable reputational damage, although firms may overreact if they devote too much attention to every minor complaint. The compliments of satisfied customers can influence others and lead to a boost in sales for the company's products, but when and how can a firm make use of such positive comments? A major task for sales and communication managers is to monitor, predict and influence this kind of mass consumer reaction and this paper goes some way to offering an alternative approach based on patterns of herding behavior. Once the patterns are understood and can be followed, further research can investigate the consequences of different marketing actions on how the herd changes from one pattern into another. Interesting marketing information is provided not just by the current herding pattern, but also by transitions between patterns. Marketers are also increasingly adapting promotion campaigns day-by-day, depending on consumers' reactions. It is not uncommon for a brand campaign to launch two slightly different versions of a campaign to test which version works best. By adding data about changes in herding patterns to this approach, marketers can not only adapt the campaign based on individual reactions or click-through rates, but also on effects on herding transitions. Campaigns can then be seen as being successful if they bring the herd into a more desirable pattern. Many companies install web teams to track online communication around their products and services and to intervene when possible. Our approach of tracking pattern transitions adds to this strategy by allowing managers to state and track the goal of a campaign more precisely. We return to this point in the discussion, to examine how our framework of herding patterns can be applied by marketing managers.

The rest of this paper is organized as follows. First we relate this study to relevant literature and describe the basis on which we build. We then propose a conceptual framework of three dimensions along which changes in mass consumer behavior are important, leading to eight different herding patterns. Following that, we describe a method for assessing the prevalence and dynamics of these patterns in the online setting, using empirical data from Twitter. We conclude with a discussion of theoretical and managerial implications and we address limitations of this study and propose avenues for further research.

Related Literature

Herding from a Marketing Perspective

There is a lack of clarity in the literature about what the term "herding" means from a marketing perspective. In an early paper on this subject, Banerjee (1992) shows herding behavior in a population of hypothetical agents who make choices between assets based on their own information and also on the observed behavior of other agents. Following this, Bikhchandani, Hirshleifer, and Welch (1998) assess convergent behavior that does not follow principles of traditional economic models and postulate that informational cascades, sanctions against defectors, network externalities, and preference effects combine to produce herding. These studies do not formally define herding but see it as a phenomenon through which people tend to converge on similar behavior, resulting in a situation with "everyone doing what everyone else is doing" (Banerjee 1992, p 798). However, this may be an oversimplified definition as it suggests that if a proportion of a population expresses the same behavior, it cannot be called a herd. It may be more useful for marketers to define herding as *a process* whereby a herd can develop in size and possibly in other ways too. Despite the subsequent attention to herding in the marketing literature, a formal definition has not emerged.

Looking to the cognitive science literature, we find a definition which suits our purpose of describing the patterns and dynamics of herding behavior. Following Rafaat, Chater, and Frith (2009, p 420) we see herding as "a form of convergent social behavior that can be broadly defined as the alignment of the thoughts or behaviors of individuals in a group (herd) through local interaction and without centralized coordination." These authors present an overview of approaches that addresses herding in humans and they distinguish between two main types of research. Pattern-based explanations of herding behavior assume that individuals can be modeled (often mathematically) as simple decision-making units and that the patterns of relationships between people are dominant in forming group behavior (Cucker and Smale 2007; Dyer et al. 2008). In contrast, transmission-based explanations assume that the complex intra-personal cognitive mechanisms play a major role in the mental processes that govern how people receive and process information and then transfer it to others (Frith and Frith 2006). In this paper we build on the first of these, pattern-based explanations of herding behavior.

Comparing Herding to Diffusion, Influence and Contagion

During the last few decades, marketing scholars have investigated mass consumer behavior using a number of related concepts. Besides herding, various studies investigate processes of diffusion, influence and contagion and it may therefore aid the reader to consider the similarities and differences between these terms so as to understand the potential that herding has for predicting influential mass behavior (see Table 1). *Diffusion* is a term stemming from the physical sciences which, when used in a marketing context, describes the spread of a product or a product-related behavior via communication channels through a social system (Bass 1969; Peres, Muller, and Mahajan 2010; Rogers 2003). The terms diffusion and herding both adopt a population-level view of mass behavior, although diffusion is focused on the product or product-related behavior that spreads through a population and herding is focused on the group itself. *Influence*, or social influence, occurs when a person's attitude or behavior is changed due to what other people say or do (Bonfield 1974). According to Cialdini (2008), influence occurs through six mechanisms: reciprocity, commitment and consistency, social proof, authority, liking, and scarcity. *Contagion*, or social contagion, is similar to influence but adopts the metaphor of viral infection such that the focus is on the behavior that makes use of people to propagate itself (Du and Kamakura, 2011; Langley et al. 2012). Influence and contagion are two processes with the focus at the individual level.

In contrast to the concepts described above, we take *Herding* to be the group process through which a population of consumers develops whereby behavior becomes unified, the number of people expressing the behavior grows, and changes spread rapidly through the herd.

Herding in the Online Setting

Although the process of herding is conceptually independent of the medium in which the herd members interact, in practical terms there are significant differences between traditional (offline) herding and herding in the online setting. According to Barry and Fulmer (2004), three important attributes of the online setting are social bandwidth, interactivity, and surveillance and we propose that these are relevant to herding as they change the way that people interact and influence each other. *Social bandwidth*, which refers to the transmission of socially-relevant information via the online medium, will affect the process of herding by providing social identity cues through which individuals can assess their similarity to others in the herd. The types of social identity cues made available online may be far more limited than when meeting others face-to-face but the cues may be highly relevant to the situation at hand, for example in an online community devoted to a specific topic. *Interactivity* may also play a role in herding in the online setting, as the exchange of information between (potential) herd members is quite different to offline interactivity. Individuals can interact with a wide range of others both real-time and asynchronously and many-to-many discussions between very large numbers of people is possible. *Surveillance*, which refers to the publicly observable nature of much online information, may also alter the way that herding behavior develops. If a small, developing herd makes its intentions visible via an online application or Web site then many other potential members can see what is happening and follow developments, as onlookers, without the need for the intervention of traditional media. Taken together, these distinctive attributes of online social media represent a foundation for a change in herding dynamics.

Patterns of Herding and its Consequences

Herding has been shown to occur in marketing-relevant behavior, such as product adoption (Hanson and Putler 1996), financial investments (Zhang and Liu 2012) and community membership (Oh and Jeon 2007). Herding in the online setting has received research attention in the marketing and broader management literature and various studies have concentrated on the effect that the perceived popularity of product choices has on the likelihood of herding behavior occurring. For example, Hanson and Putler (1996) show that consumers who download public domain and shareware software are highly influenced by the number of previous downloads of a particular program. Herding behavior develops whereby one of a pair of similar programs quickly becomes the market leader and this effect can be brought about by artificially manipulating the apparent number of downloads, thereby changing the perceived popularity. Similarly, consumers are shown to influence each other when downloading music from a web application (Salganik, Dodds, and Watts 2006). By comparing conditions where individuals download songs after listening to them versus where individuals also see the download choices of others, these authors show that social influence can drive online herding behavior and they also show that this herding produces apparently unpredictable outcomes in terms of which songs

Table 1
Comparing herding to diffusion, influence and contagion.

Behavioral phenomenon	Focus	Uses	Key literature
Herding	Population level, focus is on the group of people expressing a behavior	Understanding the development of influential consumer groups	Banerjee (1992), Zhang and Liu (2012)
Diffusion	Population level, focus is on the market penetration of a behavior	Understanding the market penetration of products	Bass (1969), Peres, Muller and Mahajan (2010), Rogers (2003)
Influence	Individual level, focus is on a person and how their behavior is changed due to what others say or do	Understanding how people can be persuaded to change their behavior	Bonfield (1974), Cialdini (2008)
Contagion	Individual level, focus is on a behavior and how it passes from person to person	Understanding how products can be designed for self-dispersion.	Du and Kamakura (2011), Langley et al. (2012)

become popular within a population. Tucker and Zhang (2011) assess online herding with wedding service Web sites, specifically focusing on the impact of popularity information on niche versus broad-appeal vendors. Their findings suggest that individuals in the herd derive quality information from popularity information: when popularity information for a niche vendor is artificially increased, it disproportionally increases herding. In another online product choice setting, Chen, Wang, and Xie (2011) investigate two social influence drivers of online herding on the book seller Web site, Amazon.com, namely word-of-mouth and observational learning. They find that herding occurs but that the two drivers have opposite asymmetric effects: negative word-of-mouth is stronger than positive, whereas positive observational learning is stronger than negative.

Herding is also shown to occur in online community membership where different social network structures can produce different effects. Oh and Jeon (2007) investigate different conditions on herding behavior for project membership in two open source software communities and they find that membership herding is particularly strong in large, scale-free network structures, such as those found on the Internet.

Recently, attention has been paid to herding behavior when bids are made in online auctions. Simonsohn and Ariely (2008) show that people bidding for products on the online auction Web site eBay often observe others' decisions before deciding themselves. They find that herding comes about as initial bids attracted to low starting price auctions cause other bidders to congregate, thereby increasing competition and actually resulting in higher prices being paid. Herzenstein, Dholakia, and Andrews (2011) show that on the online lending Web site, Prosper, lenders also exhibit herding behavior, although unlike on eBay, it appears to be beneficial to join the herd. Specifically, they find that the loan auctions that start to attract lenders bring about herding whereby more lenders make bids until the loan is fully funded. In an exploratory analysis they also show that two years on, the loan auctions which attracted more herding (i.e. more lending bids, more quickly) were more likely to be paid back on time, indicating that the process of herding may somehow make use of information regarding the trustworthiness of borrowers. Zhang and Liu (2012) also investigate lender herding on Prosper and disentangle two herding mechanisms: irrational (simple mimicking of other lenders) versus rational (active observational learning). They find that lenders engage in rational herding while making various inferences about why other lenders are attracted to auctions. Again, they show that herding is associated with fewer loan defaults, whereby rational herding has a stronger effect on loans subsequently being paid back.

In these studies, herding is measured as many people downloading the same song (Salganik, Dodds, and Watts 2006), joining the same open source software project (Oh and Jeon 2007), or offering loans (e.g. Zhang and Liu 2012). No account is taken of different patterns of herding, such as different herd sizes or herds that grow more quickly than others. To understand herding so that a firm is able to react, in real-time, to the development of a herd, we conclude that there is insufficient scientific insight into patterns of herding and how these patterns develop through time. Added to this, extant literature focuses on herding situations with a binary behavioral choice, such as lend/do not lend, buy/do not buy. In many situations related to products and brands it is interesting to understand how opinions in herds develop; for example, is there a growing consensus or is there a fragmentation of different factions?

It is too simple to imagine that there are only two herding states: that a herd either acts as one or the individuals all act independently of each other. So far there has been no assessment of progression in the development of herds, or analysis of various patterns emerging from different examples of widespread group behavior. This paper proposes such a set of herding patterns, tests empirically whether these patterns actually come about in an online setting and investigates the dynamic transitions between herding patterns.

Patterns of Herding Behavior

Herding as described in the introduction can be succinctly characterized as mass alignment without central coordination. We study mass consumer behavior, such as buying a product or posting a message online. The herd consists of all consumers showing this behavior. Within this herd there can still be some variation in the behavior. For instance, if the behavior is buying music, the variation is in the different songs users buy.

In order to study, model, understand and ultimately influence herding we now present a conceptual framework of patterns of herding behavior. Inspired by the pattern-based approach described above from the cognitive science literature (Rafaat, Chater, and Frith 2009), we propose three dimensions along which mass behavior can be described:

1. *Speed of contagion*: how fast is the behavior spreading? Do all consumers start displaying the behavior at once, or does the behavior slowly diffuse? Using a biological analogy, the process is slow when a few animals in a herd start moving and gradually more and more animals also start moving, until in the end the whole herd is on the move. The process is fast when a herd is frightened by a gunshot and starts moving all at once. In the same way a Twitter discussion on a particular topic can carry on slowly or can explode when something shocking rapidly spreads.
2. *Number of individuals*: the proportion of the whole population expressing the behavior or expressing an opinion on the topic. The magnitude is large when a relatively large number of individuals expresses the behavior and it is small when only a relatively small number of individuals from the population does so. In the Twitter example, the magnitude is large when a relatively large amount of people take part in a discussion and it is small when only a few people take part in a discussion.
3. *Uniformity of direction*: the extent to which the mass behavior is increasingly uniform with one variant of the behavior becoming increasingly dominant. For example, suppose there are two types of behavior, A and B. The herd is directed if A is consistently growing at the expense of B.

At one end of the spectrum (highly directed, or uniform) the whole herd moves as one and at the other end (highly undirected, or diverse) there is apparent chaos where all actors move individually. The Twitter discussion can either lead to a widespread consensus or there can be all kinds of different opinions.

The eight herding patterns can be described by their position on the three dimensions, above. In Table 2, examples are given of how each pattern would come about on the online application, Twitter. It can be seen that some of the patterns are not what we would normally consider to be herding, such as when the group size is small or when there is no aligned direction. Intuitively, people associate herding with the stampede pattern: a large group moving in the same direction, fast. We argue that it is of critical importance to understand patterns that lead to classical stampede herding and to be able to react in an appropriate and timely manner. Therefore, we include such 'non-herding' patterns in the herding framework in order to provide a full picture of patterns of mass behavior and to help marketers to be able to track the development of a herd through these 'low herding' patterns so that they are not surprised by a sudden, massive and aligned herd.

In this explorative study we define the following research questions. Do these herding patterns occur in practice? Is there a basic pattern that always occurs at the onset of herding behavior in the online setting? Is there a relationship between patterns, such as a regular transition of one pattern into another? In the next paragraphs we describe a method for assessing the prevalence and dynamics of the described patterns, using empirical data from the online micro-blogging application, Twitter.

Empirical Analysis in an Online Setting

In this section we attempt to find herding patterns in data from the online social networking and microblogging application, Twitter. Released in 2006, Twitter allows users to post and read small text-based messages of up to 140 characters, called 'tweets'. It has become one of the most successful Internet applications and as of June 2013 it is ranked 12th as measured by the Alexa Traffic Rank, which is a measure of activity relative to all Internet sites during the most recent three months, combining the average daily unique visitors to a site and the number of page views on the site. Twitter is relevant for researchers, marketers and other business domains as it reflects real time consumer opinion about firm-related and brand-related matters (Fischer and Reuber 2011; Jansen et al. 2009). Added to this, tweets can have real effects on marketplace outcomes (Bollen, Mao, and Zeng 2011). The behavior we study is tweeting using a given hashtag. Twitter users can include a hashtag in their tweets to signal that they are tweeting about a certain topic. For instance, they can use the hashtag #nowplaying if they are discussing which music they are listening to at that moment. Of course, there can still be variation in the subtopics discussed under a hashtag. Hashtags are a natural way to delimit herds, because people using hashtags choose to be part of the community of people tweeting on a topic. Our data set consists of "*streams*" of tweets consisting of all tweets containing a unique hashtag in a certain period. We first define measures for the three dimensions, number of individuals, speed of contagion, and uniformity of direction and analyze how these measures behave on a data set of 4,622 streams of tweets containing a certain hashtag. We classify each time period of each stream into one of the herding patterns. We then investigate which patterns, and transitions between patterns, arise in the data.

Data Collection

The streams of tweets were collected through Twitter's public application programming interfaces (APIs). For each hashtag we downloaded all tweets containing the hashtag that were available through the Twitter REST API. We defined this data set as the stream attached to this hashtag. The REST API can retrieve tweets from the last 6–9 days containing a search term up to a maximum of 1,400 tweets. Consequently, streams can have different durations, if more than 1,400 tweets were

Table 2
Conceptual examples of the eight herding patterns.

Number of individuals	Speed of contagion	Uniformity of direction	Pattern name	Twitter example
−	−	−	1. Slow meandering	Gradual spread of a topic with a range of diverse opinions, expressed by few people, e.g., many low-interest topics on Twitter
−	+	−	2. Fast meandering	Rapid spread of a topic with a range of diverse opinions, expressed by few people, e.g., argument of conflicting views in a small community via Twitter
−	−	+	3. Slow converging	Gradual spread of a dominant opinion expressed by few people
−	+	+	4. Fast converging	Rapid spread of a dominant opinion expressed by few people
+	−	−	5. Cold Brownian	Gradual spread of a topic with a range of diverse opinions, expressed by many people
+	+	−	6. Hot Brownian	Rapid spread of a topic with a range of diverse opinions, expressed by many people e.g., trending topic discussing varying opinions about a new brand campaign
+	−	+	7. Marching	Gradual spread of a dominant opinion expressed by many people, e.g., increase in interest surrounding a long-term brand-related issue
+	+	+	8. Stampeding	Rapid spread of a dominant opinion expressed by many people, e.g., trending topic about a new brand campaign which achieves a converging set of opinions

posted in the last 6–9 days. We define the start time of the stream as the time the first tweet it contains was posted.

The hashtags were selected randomly from the Twitter Spritzer API. This API returns in real-time a random 1% of all tweets being posted. We collected data from it for eight hours. No other selection criteria (e.g., on location or language) were used. In total we downloaded streams for 4,622 hashtags, containing 4,830,099 tweets in 229,791 periods of one hour.

Measures

From a preliminary investigation of a range of alternatives we have chosen the following measures for the three dimensions:

Number of individuals total number of participants in a stream;
Speed of contagion number of tweets per hour in a stream;
Uniformity of direction the convergence of subtopics in a stream.

For speed of contagion, we follow Asur et al. (2011) and use the number of tweets per hour as the measure for how fast the discussion is moving. For magnitude, we take the total number of people posting tweets with each hashtag as a way of representing the mass nature of each herd. The number of subtopics within the tweets using a hashtag says something about the direction or convergence of the opinions within the herd, and the way that we measure direction is more complex as it is a measure of the content within the tweets. To do this, we clustered the tweets containing the same hashtag into subtopics and calculated subtopic distribution using a technique called Latent Dirichlet Allocation (LDA), which is similar to latent class analysis (Blei, Ng, and Jordan 2003) and has been applied to Twitter (Ramage, Dumais, and Liebling 2010; Weng et al. 2010). We used the implementation of LDA in the Gensim package (Řehůřek and Sojka 2010). LDA determines groups of words that frequently occur together, called subtopics, and the chance that each tweet belongs to a subtopic. Using these probabilities we can calculate the distribution of subtopics. We then determined the Kullback and Leibler (1951) divergence for these subtopics:

$$D = \sum p_i \log\left(\frac{p_i}{q_i}\right)$$

where p_i are the probabilities in the distribution per hour and q_i are the probabilities in the cumulative distribution for all previous hours. The Kullback–Leibler divergence is a measure for how predictable the tweets in the current hour are based on the previous hours.

The correlations between measures for the three dimensions are low (.33 between those for number of individuals and speed of contagion, −.21 between those for number of individuals and uniformity of direction), except for the correlation between speed of contagion and uniformity of direction (−.81) which is as expected; generally, the faster a topic spreads from person-to-person, the less directed or less predictable the tweets are. However, it is the rare events when a mass of people develop a united opinion that are of particular interest to scientists, managers and policy makers. The frequency of herding patterns in the data is as shown in Table 3.

All eight of the herding patterns proposed in the conceptual framework are found to occur in the data set. Slow meandering is the most frequent pattern, because all hours in which no tweets are posted fall under this pattern (33% of all hours in our data set). Given that our data set excluded streams with fewer than 200 tweets in eight hours, the true total of this pattern will be much higher. However, most streams begin in the Fast meandering pattern or even in the Hot Brownian pattern, showing how quickly Twitter streams can develop.

To illustrate the choice of measures we look at one example stream: all tweets containing the hashtag #beliebershelpbeliebers. Fans of the teenage idol Justin Bieber use this hashtag to ask each other for help (e.g., to find more followers or to make a Bieber-related topic trending). As described above, the uniformity of direction dimension measures the distribution of tweets over the different subtopics in the stream. The subtopics are determined by LDA, by clustering tweets based on the words they have in common. The ten subtopics for this stream are listed in Table 4. Initially, the first subtopic was most common but there was a wide diversity of other subtopics in the stream. As the herd grew, the uniformity of direction remained low and the herd developed from Fast meandering to Hot Brownian. At a certain point, the second subtopic became very popular and the tweets converged on this while the other subtopics diminished, resulting in a uniform direction whereby the stampeding pattern emerged.

For each stream we measured the sequence of herding patterns in consecutive periods. The empirical transition probabilities between the patterns in such sequences are as shown in Table 5.

It is interesting to notice that for five patterns (Slow and Fast meandering, Slow converging, Hot Brownian and Marching), the transition probabilities indicate that it is most likely that the same pattern will also occur in the next period. One might say that these patterns are stable over time. The transition probabilities indicate that the Slow meandering pattern is the most stable over time (with a transition probability to itself of 91.9%). In contrast, three patterns are inherently unstable (Fast converging, Cold Brownian and Stampeding) because the transition probabilities indicate that

Table 3
Frequency of occurrence of the eight herding patterns in the twitter data.

Herding pattern	Slow meandering	Fast meandering	Slow converging	Fast converging	Cold Brownian	Hot Brownian	Marching	Stampeding
Total number of periods in this pattern	75,242	26,791	43,443	3,655	4,973	42,495	27,137	6,055
Percentage of periods in this pattern	33%	12%	19%	2%	2%	18%	12%	3%
Number of topics that start in this pattern	556	2,334	541	20	2	947	1	4
Percentage of topics that start in this pattern	13%	53%	12%	0%	0%	21%	0%	0%

Table 4
Ten subtopics in the stream #beliebershelpbeliebers.

Topic	Words	Example tweet
1	plzzz, vote, right, http, favor, esto, rt, por	@– VOTE 4 ME PLZZZ #BELIEBERSHELPBELIEBERS right ?? http://t.co/KQ1Wwxvw
2	buy, let, mistletoe, help, rt, thanks, ema, trip, voting	@– could you please help trend BIEBERBLASTING MISTLETOE and spread the word thanks:) #beliebershelpbeliebers â™¥
3	everyone, biebs, venezuela, welcometovenezuelambieber, someone, please, retweet	WelcomeToVenezuelaTeamBieber justin se sorprendera y nunca olvidara a venezuela! :$ #BeliebersHelpBeliebers
4	f4f, belieber, thanks, problem, ema, mtv, voting, trip, plz	@– PLZ HELP @– WIN A TRIP TO THE MTV EMA'S BY VOTING HERE – THANKS! #BeliebersHelpBeliebers
5	made, twitter, xxxx, noticesbiebs, back, hey, please, follow, rt, claus	Ayuden a #noticesbiebs a que BIEBER CLAUS se vuelva TT:D â™¥ #BeliebersHelpBeliebers
6	followback, lo, portugal, help, thanks, beliebs, heey, follower, mt, win	RT @–: PORTUGAL BELIEBS #BeliebersHelpBeliebers (:
7	followme, followers, lookatmenow, problem, rt, muchlove, following, xxx, get	Everyone Help Me To Get More Followers !! #BeliebersHelpBeliebers #LookAtMeNow #FollowMe
8	sure, lovethejdbiebs, meet, amazing, mybelieberside, nada, dont, song, justin, page	@– You should follow @– ! She's amazing! And she deserves to meet Justin! #beliebershelpbeliebers #Swag ^_^
9	retweet, beliebers, beliebershelpbeliebers, help, http, following, follow, rt, justin, welcome	@– you're welcome, #BeliebersHelpBeliebers :)
10	sing, please, http, leggoswagg	@– Please sign? #BeliebersHelpBeliebers http://t.co/SrAzOI1w

Note: The column 'Words' gives typical words per subtopic, as determined by the LDA method (see text). The last column gives an example per subtopic. For privacy reasons users names have been replaced by '–'.

it is less likely that the same pattern will also occur in the next period. A Cold Brownian pattern, for example, is most likely to move into a Marching pattern (47.7%) in the next period. These three patterns indicate that the herd is in a state of transition. Because our measure for magnitude, the total number of participants, can only increase, there can be no transitions from the group patterns (Cold and Hot Brownian Motion, Marching and Stampeding) towards the individual patterns (Meandering and Converging). This is a limitation of the current implementation. All other transitions can occur. Importantly, streams usually progress from the individual patterns to the group patterns via one of the fast individual patterns: Fast meandering or Fast converging. Streams can go directly from Fast converging to Stampeding. But if they are in Fast meandering they are vastly more likely to first go through Hot Brownian (see Fig. 2). This means that streams that end up Stampeding typically follow a standard dynamic: they increase first on the speed of contagion dimension, then increase on the number of individuals dimension before finally becoming more uniformly directed. It is of scientific and practical importance to further study such progressions so that we are able to reliably predict when a topic will represent an important mass behavior or opinion.

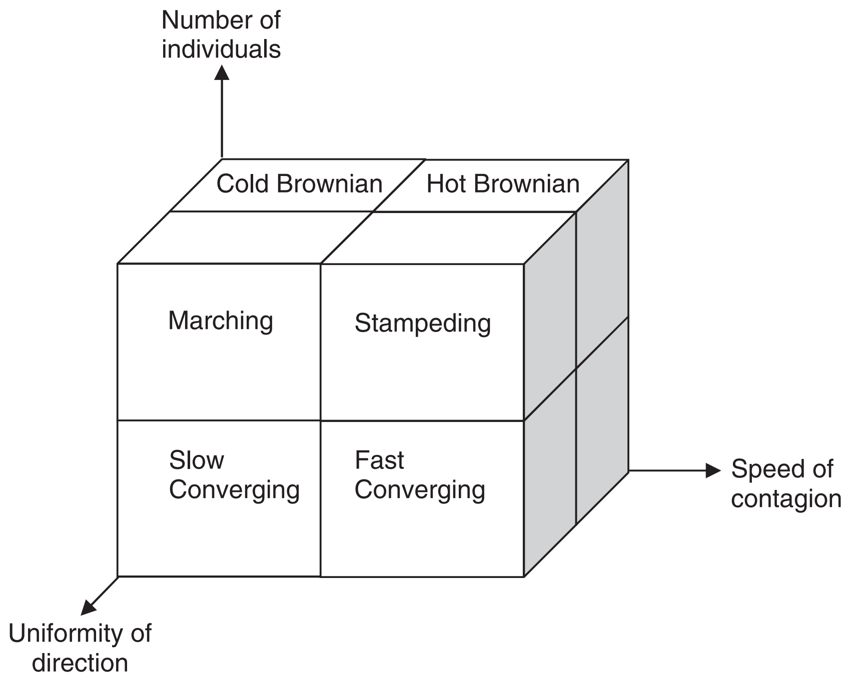
Discussion

In this paper, we take the first steps in forming a basic structured framework for distinguishing between different types of herding. Different cases of herding have different characteristics; some spread more quickly, some reach more people and in some a consensus emerges. Some of these characteristics may make one herd important for firms or governments whereas others require less attention. The first objective of this paper is to propose a framework, based on the pattern-based explanation of herding from the cognitive science literature, by which we can identify

Table 5
Transition probabilities between the eight herding patterns.

		To							
		Slow meandering	Fast meandering	Slow converging	Fast converging	Cold Brownian	Hot Brownian	Marching	Stampeding
From	Slow meandering	91.9%	2.5%	5.2%	.3%	.0%	.1%	.0%	.0%
	Fast meandering	6.7%	69.8%	10.9%	4.8%	.1%	7.5%	.1%	.2%
	Slow converging	25.2%	5.4%	66.3%	2.5%	.0%	.2%	.3%	.1%
	Fast converging	10.1%	31.4%	34.8%	20.7%	.1%	1.4%	.3%	1.1%
	Cold Brownian					17.4%	27.0%	47.7%	7.9%
	Hot Brownian					4.0%	82.3%	7.8%	5.8%
	Marching					8.2%	9.5%	76.6%	5.7%
	Stampeding					5.9%	36.9%	28.4%	28.8%

Note: When a Twitter stream exhibits the slow meandering pattern, there is a 91.9% chance that it will remain so in the next hour, a 5.2% chance that it will develop into the slow converging pattern, and so on.



Note: The two patterns not visible are Slow Meandering (back, lower left) and Fast Meandering (back, lower right). The Marching and Stampeding patterns are perhaps the only 'true' herding patterns in the generally accepted sense of the word, although this framework aids our understanding of the related patterns which lead to them.

Fig. 1. A conceptual framework of patterns of mass consumer behavior (herding).

different patterns of mass behavior. We describe a conceptual framework whereby eight different patterns can be distinguished based on three dimensions: the *number of individuals* expressing a behavior, the *speed of contagion* with which the behavior spreads through a population, and whether there is a *uniformity of direction* to the behavior. The eight resulting herding patterns cover a wide variety of herding types from small, indistinct clusters through to rapid, large-scale and united crowds.

Our second objective is to test the applicability of the framework, by assessing which of these patterns can be witnessed in practice using a large database of online behavior taken from the popular micro-blogging Internet application, Twitter. We explore different possible metrics for each of the dimensions and then measure the position of streams of tweets on each dimension. From this, we are able to identify not only the prevalence of the eight herding patterns, but also to follow the progression of each herd as it grows, develops and disperses. It is interesting to note that in our data, it is relatively rare that herds exhibit both a high speed, many tweets per hour on a subject, and a converging direction. This seems logical but it is precisely these rare events that may be of particular interest to managers, as it is these herds that most likely require their attention. The proposed framework allows for the easy identification of such noteworthy herds.

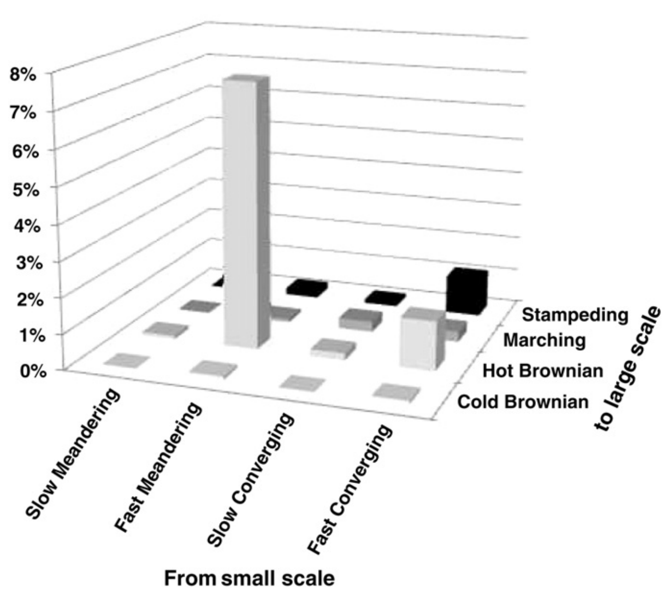


Fig. 2. Transitions between herding patterns in twitter data, from small scale patterns to large scale patterns.

Implications

Scientifically, the herding framework presented in this paper may help with theory building for herding behavior. Future theories or models of herding behavior may take the different herding patterns into account because they help to separate the important group opinions from the background noise. Until now, herding theory development has focused mainly on the types of social interactions that can lead to a herd emerging, such as mimicry or observational learning (Banerjee 1992;

Bikhchandani, Hirshleifer, and Welch 1998; Chen, Wang, and Xie 2011). No attention has yet been paid to mechanisms driving behavioral alignment at different stages in a herd's development, and yet it is quite likely that different processes of social influence drive an initial, small herd, for example in the fast meandering pattern, than a later, large herd. Our proposed framework may help researchers to identify phases of herd development.

Added to this, our results show that, in the online setting we study, there is a dominant path of herd development: most herds that become a stampede follow a consistent transition path as they first increase on the speed of contagion dimension, then on the number of individuals dimension and then finally on the uniformity of direction dimension. This suggests a sequence of different social influence processes, and it would be useful to understand why the transition probabilities shown in Table 5 and Fig. 2 occur. For example, if our results are shown to be robust, why do herds that begin small and that initially converge to a uniform direction, almost never develop into large-scale stampedes? Theories that can help to explain the dominant transition path we find may enrich the herding patterns framework; perhaps by showing that herds that are not too uniform tend to be inclusive, whereby many potential participants feel that their opinion fits, allowing the herd to grow in number before a dominant opinion emerges.

Managerially, this paper provides a structured set of herding types, which is a necessary step in allowing firms to understand more clearly on a broad scale what their consumers think and what they are doing. By following real-time changes in the herding patterns related to a brand campaign, marketing managers can adapt the timing of their actions. The transition probabilities between herding patterns provide guidance for this type of conclusion that may be investigated further. For instance, if they want to stimulate a stampede, and they observe an increase in the number of individuals participating in the discussion, whereby the transition takes place from Fast meandering to Hot Brownian, then they should focus their efforts on increasing the uniformity of direction. Indeed, the finding that there is a dominant transition path between herding patterns has a strong implication: if this result is found to be robust across a variety of data sets, then any organization that is particularly interested in identifying a stampede needs to track only emerging herds following this sequence (whereby speed, then magnitude, then direction increases), thereby reducing the need to spend resources following online interaction for a vast number of topics. And any organization interested in creating a stampede can take appropriate action to develop the herd along this development path.

Managers can choose their marketing actions based on the state a herd is in, and they can intervene before a stampede emerges. For instance, we speculate that it may be difficult to change the direction of a herd once it has reached the Marching or Stampeding patterns, whereas a well-placed intervention could tip the balance in a herd in the Hot Brownian state. Similarly, a herd that is in the Fast meandering pattern (small and non-uniform direction, but diffusing quickly) may be a threat in the long term if it expresses negative sentiments to a brand, whereas a herd in the Slow converging pattern (small and slowly spreading, but with a converging direction) is likely to remain a small herd that will disappear. This can be derived from the transition probabilities in Table 5.

The subsequent question for managers is *how* to bring about a desired shift in the herding pattern. This study does not investigate different intervention strategies although the three dimensions of the herding framework provide some indication of how pattern transitions can be stimulated. For example, managers can intervene by generating more attention for a brand-related issue (measured by number of individuals), but also by more subtle effects like increasing susceptibility by priming (measured by speed of contagion), or influencing people who have already joined the herd to accept a particular viewpoint related to one of the subtopics (measured by uniformity of direction).

Limitations and Further Research

The conceptual framework shown in Fig. 1 is the first to provide a structure for understanding different types of herding patterns. However, this paper has only begun to investigate such patterns. There are alternative versions of our framework possible and an interesting avenue for further research would be a comparison of alternative dimensions and alternative patterns. Some options are exploring patterns like the long tail with some dominant and many minority camps, growing discord or dissent in the population, etc.

Obviously, besides contagion through the herd, external influences will also play a role in online opinion dynamics, one example being *The New York Times* that publishes a weekly roundup of trending topics on twitter that are related to New York. This exogenous effect is important and has been shown to be a driver of behavior (Berger and Milkman 2012). However, our objective is not to explain underlying mechanisms of online mass behavior but to propose a new way of describing such behavior, providing added guidance for the type and timing of interventions. A potential future model could add to the herding patterns by including endogenous and exogenous influences on the transition probabilities.

A limitation of our empirical study is that Twitter is mostly used for sharing up-to-the-minute news items, gossip and opinions relevant to events of that day. A next step in our research is to investigate herding patterns with data from other online applications, such as social networking sites (e.g. Facebook), online discussion forums (e.g. Huffington Post), and content communities (e.g. YouTube). Besides this, our empirical study made the assumption that hashtags on Twitter are a good way of identifying herds. However, some hashtags may be very narrowly defined and others more broadly defined. It could be that a number of narrowly defined hashtags actually belongs to a single, undirected, herd. We assume for this study that this is not a significant problem, as people use hashtags for the very purpose of connecting with others interested in the same topic. The natural process of discussion will mean that dominant hashtags will quickly appear. Further research, including data from other online sources, will determine the robustness of the herding framework and our suggested metrics.

The objective of this paper is to identify different forms of herding, as online mass behavior that becomes aligned is not a binary phenomenon. However, once we become better able to distinguish between herding patterns a highly relevant question becomes, What are the effects of the different patterns on marketing indicators, such as sales or stock prices? Investigating such marketing outcomes to show the relevance of different herding patterns, or certain dynamic changes between patterns, is an important area for future research.

We conclude by mentioning a promising field of research that can be applied to the herding phenomenon: multi-agent simulation. The herding patterns discussed in this paper describe characteristics of behavior at the population level. We have argued that awareness of the population's behavior is useful for firms or governments in the context of marketing, policy support and public opinion. However, the impact of interventions by these organizations is measured via the behavior of individuals. Therefore, an important next step in this research is to understand the mechanisms in individual behavior that result in population level herding patterns as emergent consequences. Agent-based modeling (ABM) is a modeling technique that is highly suited to connect micro level behavior to macro level phenomena.

Online mass behavior that is aligned but not centrally coordinated, which we call herding, is an important phenomenon that requires more thorough scientific investigation. This will provide those confronted with herding in practice with instruments to be able to intervene efficiently.

Acknowledgement

The authors wish to thank the editor, Charles Hofacker, and the anonymous reviewers for their insightful comments and patient guidance in helping to improve the earlier version of this paper.

References

Asur, Sitaram, Bernardo A. Huberman, Gabor Szabo, and Chunyan Wang (2011), *Trends in Social Media: Persistence and Decay*. Arxiv preprint 102.1402).

Banerjee, Abhijit V. (1992), "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 3, 797–817.

Barry, Bruce and Ingrid Smithey Fulmer (2004), "The Medium and the Message: The Adaptive Use of Communication Media in Dyadic Influence," *Academy of Management Review*, 29, 2, 272–92.

Bass, Frank M. (1969), "A New Product Growth Model for Consumer Durables," *Management Science*, 36, 9, 1057–79.

Berger, Jonah and Katherine L. Milkman (2012), "What Makes Online Content Viral?," *Journal of Marketing Research*, 49, 2, 192–205.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1998), "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *The Journal of Economic Perspectives*, 12, 3, 151–70.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993–1022.

Bollen, Jonah, Huina Mao, and Xiaojun Zeng (2011), "Twitter Mood Predicts the Stock Market," *Journal of Computational Science*, 2, 1, 1–8.

Bonfield, E.H. (1974), "Attitude, Social Influence, Personal Norm, and Intention Interactions as Related to Brand Purchase Behavior," *Journal of Marketing Research*, 11, 4, 379–89.

Chen, Yubo, Qi Wang, and Jinhong Xie (2011), "Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational Learning," *Journal of Marketing Research*, 48, 2, 238–54.

Cialdini, Robert B. (2008), *Influence: Science and Practice*. 5th edition. Boston: Pearson.

Cucker, Felipe and Steve Smale (2007), "On the Mathematics of Emergence," *Japanese Journal of Mathematics*, 2, 197–227.

Du, Rex Yuxing and Wagner A. Kamakura (2011), "Measuring Contagion in the Diffusion of Consumer Packaged Goods," *Journal of Marketing Research*, 48, 1, 28–47.

Dyer, John R.G., Christos C. Ioannou, Lesley J. Morrell, Darren P. Croft, Iain D. Couzin, Dean A. Waters, and Jens Krause (2008), "Consensus Decision Making in Human Crowds," *Animal Behaviour*, 75, 2, 461–70.

Fischer, Eileen and A. Rebecca Reuber (2011), "Social Interaction via New Social Media: (How) Can Interactions on Twitter Affect Effectual Thinking and Behavior?," *Journal of Business Venturing*, 26, 1, 1–18.

Frith, Chris D. and Uta Frith (2006), "The Neural Basis of Mentalizing," *Neuron*, 50, 4, 531–4.

Gu, Chao (2011), "Herding and Bank Runs," *Journal of Economic Theory*, 146, 1, 163–88.

Hanson, Ward and Daniel Putler (1996), "Hits and Misses: Herd Behavior and Online Product Popularity," *Marketing Letters*, 7, 4, 297–305.

Herzenstein, Michal, Utpal M. Dholakia, and Rick L. Andrews (2011), "Strategic Herding Behavior in Peer-to-Peer Loan Auctions," *Journal of Interactive Marketing*, 25, 1, 27–36.

Jansen, Bernard J., Mimi Zhang, Kate Sobel, and Abdur Chowdury (2009), "Twitter Power: Tweets as Electronic Word of Mouth," *Journal of the American Society for Information Science and Technology*, 60, 11, 2169–88.

Jun, Seung-Pyo (2012), "A Comparative Study of Hype Cycles among Actors Within the Socio-technical System: With a Focus on the Case Study of Hybrid Cars," *Technological Forecasting and Social Change*, 79, 8, 1413–30.

Kullback, Solomon and Richard A. Leibler (1951), "On Information and Sufficiency," *Annals of Mathematical Statistics*, 22, 1, 79–86.

Langley, David J., Tammo H.A. Bijmolt, J. Roland Ortt, and Nico Pals (2012), "Determinants of Social Contagion during New Product Adoption," *Journal of Product Innovation Management*, 29, 4, 623–38.

Oh, Wonseok and Sangyong Jeon (2007), "Membership Herding and Network Stability in the Open Source Community: The Ising Perspective," *Management Science*, 53, 7, 1086–101.

Peres, Renana, Eitan Muller, and Vijay Mahajan (2010), "Innovation Diffusion and New Product Growth Models: A Critical Review and Research Directions," *International Journal of Research in Marketing*, 27, 2, 91–106.

Rafaat, Ramsey M., Nick Chater, and Chris Frith (2009), "Herding in Humans," *Trends in Cognitive Sciences*, 13, 10, 420–8.

Ramage, Daniel, Susan Dumais, and Dan Liebling (2010), "Characterizing Microblogs with Topic Models," *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, May 23–26, Washington, DC.

Řehůřek, Radim and Petr Sojka (2010), "Software Framework for Topic Modelling with Large Corpora," *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 19–21 May, Valletta, Malta. p. 45–50.

Rogers, Everett M. (2003), *Diffusion of Innovations*. 5th edition. New York: Simon & Schuster.

Salganik, Matthew J., Peter Sheridan Dodds, and Duncan J. Watts (2006), "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market," *Science*, 311, 5762, 854–6.

Simonsohn, Uri and Dan Ariely (2008), "When Rational Sellers Face Nonrational Buyers: Evidence from Herding on eBay," *Management Science*, 54, 9, 1624–37.

Tucker, Catherine and Juanjuan Zhang (2011), "How Does Popularity Information Affect Choices? A Field Experiment," *Management Science*, 57, 5, 828–42.

Weng, Jianshu, Ee-Peng Lim, Jing Jiang, and Qi He (2010), "TwitterRank: Finding Topic-sensitive Influential Twitterers," *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, 3–6 February, New York, NY.

Zhang, Juanjuan and Peng Liu (2012), "Rational Herding in Microloan Markets," *Management Science*, 58, 5, 892–912.