Predicting Emerging Social Conventions in Online Social Networks

Farshad Kooti
MPI-SWS
farshad@mpi-sws.org

Winter A. Mason
Stevens Institute of Technology
m@winteram.com

Krishna P. Gummadi
MPI-SWS
gummadi@mpi-sws.org

Meeyoung Cha
KAIST
meeyoungcha@kaist.edu

ABSTRACT

The way in which social conventions emerge in communities has been of interest to social scientists for decades. Here we report on the emergence of a particular social convention on Twitter—the way to indicate a tweet is being reposted and attributing the content to its source. Despite being invented at different times and having different adoption rates, only two variations became widely adopted. In this paper we describe this process in detail, highlighting the factors that come into play in deciding which variation individuals will adopt. Our classification analysis demonstrates that the date of adoption and the number of exposures are particularly important in the adoption process, while personal features (such as the number of followers and join date) and the number of adopter friends have less discriminative power in predicting adoptions. We discuss implications of these findings in the design of future Web applications and services.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences; H.2.8 [**Database Applications**]: Database Management—*data mining*

Keywords

Social Conventions, Microblog, Prediction

1. INTRODUCTION

The importance of social norms in a society can hardly be understated. They can affect our most basic perceptions [2] and are the basis of massive social movements [34]. They guide everyday behaviors of the general public [1] as well as those of the most extreme members of a society [43]. As a consequence of their importance, social norms have been the topic of a multitude of studies in anthropology [5], sociology [24], and psychology [17, 41].

Social norms may begin as social conventions, i.e., simple habits of social interaction. Over time, some social conventions become more strongly associated with the group identity and some even become so central that failure to obey the social convention results in censure—at which point they have transitioned from a social convention to a social norm. This transformation can serve multiple purposes; it reinforces the system that acts as a guide in social interactions [21, 33], enhances group identity and feelings of belongingness [36], and facilitates group performance on tasks by establishing standardized roles [30].

The factors influencing an individual's decision to conform to a social norm have been studied extensively in psychology [17]. In some cases, the individual is adopting a norm because they believe doing so leads to more accurate and useful judgments; that is, because they believe it is practically useful to adopt the norm [20] and their peers are providing "social proof" [16]. For instance, one may believe that the reason everyone owns an iPod is because they are objectively better than the alternative portable music players. In these cases, factors like unanimity [2], the publicness of the action [20], the authority or personal importance of the source(s) [9], and the size of the relevant group [9, 26] all affect the probability that an individual will adopt a social norm.

In other cases, the motivation to adopt a social norm has more to do with a desire to build and maintain relationships with group members. In this case, the decision to adopt a norm has less to do with the correctness of the normative behavior but instead more with "fitting in." As before, the publicness of the action [28] is important, as is the extent to which the individual identifies with the group that embodies the norm. Similarly, the extent to which the norm is enforced by other members of the group [32] is important in the decision to adopt it. There are also individual differences that affect the likelihood of adopting a social norm, such as one's need for social approval [47].

However, nearly all of these conclusions are based on the assumption that a social norm is firmly established in the group before individuals are in a situation in which they could be influenced by the norm. While the factors that lead an individual to conform to a social norm of a group are well-studied, the factors that come into play when an individual decides to adopt a *social convention* are not as well understood. In this case, the decision to adopt a particular convention has an impact on whether that convention comes to be widely adopted. This in turn could affect the likelihood it will ever become associated with group identity or become proscriptively normative.

Social conventions can be established in many different ways. They may be created and enforced by institutions, such as the International System of Units for units of measurements. Many times, however, there is no established institutional code or even an in-

CIKM'12, October 29–November 2, 2012, Maui, HI, USA.

stitution, in which case a convention *evolves*, arising from interpersonal interaction and spreading from person to person within a community [33].

In this research, we observe the emergence of a social convention that evolves over time and document the characteristics of the convention, the environment, and the dynamics of competing variations of the convention. Specifically we focus on the Twitter online community and the convention used to indicate that one is reposting a tweet to one's followers while attributing the content to the source—an act commonly known as "retweeting" [10]. We study seven variations of this convention: RT, via, Retweeting, Retweet, HT, R/T, and the recycle icon. Utilizing a near-complete collection of tweets, we track the birth and spread of each of these retweeting variations over the first 3.5 years of Twitter's existence.

There are several unique advantages in studying the retweeting convention. To begin with, the convention is specific to Twitter, therefore is unlikely to be imported from existing conventions in the way other conventions have been, such as greetings in other online communities [7]. This also means that before the convention was established, a user was most likely to encounter it through his or her contacts on Twitter. Consequently, another advantage of studying retweeting is that one's contacts on Twitter are explicit, so we can observe how a convention spreads from one user to another. A third advantage is that the convention is not usable outside of Twitter, so nearly all uses of the convention will have been on Twitter, limiting potential exposures outside the environment.

Because we have nearly all of the tweets, this means that we also have nearly all of the uses of the convention. In other words, by studying this particular convention we can explore not just the micro-level processes of adoption or the macro-level outcomes but actually explore how one led to the other. Among many findings, some of our key results are as follows:

1. The final reach of the retweeting variations does not seem to be strongly related to either the amount of time each variation had to grow or the rate at which it initially grew.

2. Nearly all adopters had been exposed to the convention through their friends on Twitter, suggesting that social relations play an important role in the adoption process.

3. When users adopted multiple variations, they were loyal to certain variations over others. The longer variations such as Retweeting had higher probabilities of being abandoned.

4. RT and via were the most popular in almost all geographical regions, indicating that a common retweeting practice was adopted independent of cultural and linguistic borders.

5. The decision to adopt a particular variation has more to do with the global popularity of the variation rather than its popularity in the local neighborhood or personal preference.

6. The number of exposures had more discriminative power than the number of adopter friends in predicting adoptions, suggesting that volume may be more important than independent sources—contrary to other social diffusion models.

We built a classifier that predicts which variation individuals would have adopted at a given time—*micro-level prediction*. Despite much effort, it was increasingly clear from our analysis that the micro-level processes of convention adoption could not be easily predicted. Before adopting a variation, most users (96.71%) had been exposed to it at least once, but this was not sufficient to predict the adoption. Considering a wide range of features, our classifier could achieve an accuracy of on average 22% improvement against the baseline (i.e, predicting 61% of cases correctly) and up to 65% improvement for less popular variations (i.e., predicting 82.3% of cases correctly). After ensuring stronger definitions of social ties (i.e., mentioning links) and adoptions (i.e., usage of at least 3 or 5 times), the overall prediction accuracy improved by an additional 10%.

The uncertainty in the outcome of which variation will be adopted at the micro-level perhaps may be intrinsic to the process of convention spreading. Similar to the uncertainty introduced by social feedback mechanisms [39], the choice of adoption may be inherently difficult to predict reliably or accurately. This is especially likely given that a similar uncertainty is observed at the macro-level; what was once the dominant variation (e.g., via, Retweeting) was no longer the most popular after a few months.

This paper is one of the first to analyze the underlying process in the establishment of a social convention with a massive data set. We believe our research, although preliminary, opens the door to new methods for studying social conventions. In the future, we hope to replicate this study over a much longer period of time (e.g., 10 years) to examine how social conventions turn into social norms.

The remainder of this paper is organized as follows. In Section 2, we review literature on social norms, social conventions, and social influence. In Section 3, we describe the Twitter data set and our methodology to extract variations of the convention. Section 4 presents the analysis of possible features that could be used to determine the adoption process of the convention and Section 5 presents the results of the prediction analysis based on these features. In Section 6, we discuss further implications of our work and we conclude in Section 7.

2. RELATED WORK

One of the first studies that observed the effect of norms on human behavior was Sherif's study on the *phi* effect [41]. In his study, peoples' estimates of how much a dim light moved in a dark room not only stabilized when participants were organized in pairs or teams, but also converged to each other's estimates. The influence of other people's judgments affected the participants' judgments of simple perceptual phenomena. This effect was more dramatically demonstrated by Asch [2], who asked participants to decide which of three lines was the shortest. One line was clearly shorter, but when a set of confederates all chose a longer line, participants conformed to the group's decision and gave an incorrect response in 37% of cases.

Since these studies established the powerful effects social norms can have, others have sought to identify the features that influence the decision to adopt a social norm. Cialdini and Trost [17] review a large body of work and highlight a few of the factors that have been identified as most important. One of the main motivations they cite is merely the accurate interpretation of the world; when many peers adopt a particular behavior, it is a signal that the behavior is optimal in some way [20]. For instance, Stanley Milgram and colleagues found that a large number of people were induced to stare at blank space simply by having others engage in the same behavior [31]. The other main motivation Cialdini and Trost identify is the need to belong [6]. When this social affiliation need is activated, the perceived *injunctive* norm—what one's peers approve or disapprove of—becomes more important than the *descriptive* norms—what one's peers actually do. These injunctive norms can be powerful; one study found it to be the second largest predictor of whether one would engage in extradyadic sex [12].

For both kinds of motivations, certain factors can influence the likelihood of adopting the norm. The perceived relative proportion [9] and absolute number [26, 28] of one's peers who have adopted the norm are positively related to the likelihood of adoption. The greater the perceived similarity between one's self and the perceived group, the more likely one is to adopt the norm [22]. Similarly, the more one values the group [9,25,32], the more likely one will be to adopt the norm. However, most or all of these factors in the decision to adopt a social norm assume that the norm is firmly established in the group. There has been much less work on how people respond when social conventions are just emerging.

There have been laboratory studies on how conventions arise, most of which frame the emergence of a social convention as a solution to a coordination problem. One of the first to explicitly investigate how people solve coordination problems focused on linguistic conventions for reference. In [46], participants were asked to develop short-hand verbal code for referring to images that allowed them to communicate more efficiently and therefore complete tasks with the images more quickly. As a result of the need to coordinate efficiently, the participants developed linguistic or pseudo-linguistic conventions. Such an experiment in a controlled environment allows the examination of the exact steps involved in the process of creating a social convention.

A more popular approach to study the emergence of social conventions has been to create mathematical or computational models that embody a theory, in order to demonstrate how a specific mechanism could lead to the emergence of social conventions [3, 11]. In many of these models, the decision of agents to adopt a norm or a convention is assumed in the model [44] or proves beneficial to the agents because of the reward structure that is implicit in the model [27,42]. Even models that factor in social network structures in the emergence of a social norm do not focus on the individual's decision to adopt the corresponding social convention [19,40].

It seems plausible that when social conventions emerge, the individual's decision to adopt a particular variation of a convention may not be simply the economically rational outcome of a coordination problem, but instead may be influenced by other factors such as the number or proportion of one's peers who are using that particular variation and the desire to affiliate with the group that is using the variation. Moreover, since peer influence is clearly a factor in the decision to adopt an established norm and in the resolution of coordination problems, it is probable that the process may resemble other forms of social diffusion.

In one of the seminal works on social diffusion, Christakis and Fowler find evidence for the spread of obesity [15] and cooperative behavior [23] through social networks (but see [18]). In [13], Centola studies the spread of behavior in an online social network experiment with almost 1500 users. The study shows that individual adoption is much more likely when participants received social reinforcement from multiple neighbors in the social network. Even more importantly, Bakshy and colleagues [4] conducted an experiment on Facebook to demonstrate the role of social networks in information diffusion. They randomly assigned users to be exposed to a link shared by a friend or to have the link hidden and found that users who were exposed to peer influence are 7 times more likely to share a piece of information than those who never saw the link.

These studies strongly suggest that peer influence is critical in the spread of information and normative behavior [15]. In this work, we seek to find the factors that influence an individual's decision to adopt one variation of a social convention over another. The factors we consider include the behavior of one's peers, as well as the number and proportion of one's peers who have adopted it. By identifying which factors play a key role, we hope to gain a deeper understanding of the micro-level process underlying the emergence of a social convention.

3. METHODOLOGY

This paper focuses on the retweeting convention in Twitter, which is used for indicating when a message is reposted while attributing the source. In this section, we describe the Twitter dataset and its high-level characteristics. We also present the methods we employed to identify the retweeting convention and briefly introduce the seven retweeting variations that we analyze in this paper.

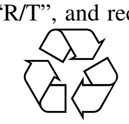
3.1 Dataset

We obtained the Twitter data set from [14], which comprises the following three types of information: profiles of 54 million users, 1.9 billion directed follow links among these users, and 1.7 billion public tweets posted by the collected users. The oldest tweet in the data set is from March 2006, when the Twitter service was publicly launched. The follow link information is a snapshot taken at the time of data collection in September 2009, while the user and tweet information is near-complete because user IDs were sequentially queried from all possible ranges (0–80 million) at the time of data collection. Therefore, this data set provides a unique opportunity to study the birth and progress of new collective behaviors in Twitter.

The Twitter network exhibits topological features that distinguish it from other online social networks. First, although the distributions for both in- and out-degree are heavy-tailed similar to other networks, certain Twitter users have extremely large in-degrees (i.e., the number of followers), at a scale that is unprecedented. The maximum in-degree is on the order of millions (as opposed to tens of thousands in other social networks). This is because Twitter users include celebrity and mainstream media accounts like Oprah Winfrey and BBC. Second, the social links are directional and only 23.8% of the links are reciprocal. In contrast, in other social networks social links are bidirectional by design or the majority are reciprocal. These differences should be taken into account when trying to apply the results of convention propagation described in this paper to a broader context.

3.2 The Retweeting Convention

Several different variations of the retweeting convention emerged during the first few years of Twitter [29]. They arose organically and became widely adopted by many individuals and third-party applications, until and even after Twitter rolled out the official, built-in "retweet" button in November 2009. The retweeting convention typically had a syntax of "token @username repeated-text" [10]. We searched for this syntax in the tweet data set to find all potential variations of this convention. Among them, we study the four most frequently used variations ("RT", "via", "Retweet", and "Retweeting") and three lesser used ones ("HT", "R/T", and recycle icon[1], shown below).



RT is short for "retweet" and HT stands for "heard through" or "hat tip."

We say a user *adopted* a given retweeting variation if the user has employed the variation at least once. Table 1 shows the number of adopters and tweets according to this definition. The two most popular variations, RT and via, reach 1.8 million and 750 thousand adopters respectively, yet the rest reach a far fewer number of adopters. In total, 2 million or 3.7% of all Twitter users adopted the retweeting convention and an impressive 59 million or 3.5% of all tweets ever posted used any variation of this convention.

[1] Can be used by copying and pasting the symbol from Unicode symbol map.

convention	# of adopters	# of usages	First usage
RT	1,836,852	53,221,529	2008-01
via	751,547	5,367,304	2007-03
Retweeting	50,400	296,608	2008-01
Retweet	36,601	110,616	2007-11
HT	8,346	22,657	2007-10
R/T	5,300	28,658	2008-06
recycle icon	3,305	18,255	2008-09
Total	2,059,350	59,065,627	-

Table 1: The number of adopters and tweets per convention

The table also shows the date when the seven variations first appeared. The birth of the retweeting convention establishes that there is a need on Twitter to indicate a message is passed on from another source and to attribute the message to the source. The first variation ever used was via in the following context,

> @JasonCalacanis (via @kosso) - new Nokia N-Series phones will do Flash, Video and YouTube

which indicates that the original tweet message came from another user (@kosso). This tweet was posted in March of 2007, only twelve months from the launch of Twitter and only 4 months after the first "@username" reference appeared in Twitter.

For most variations, their invention seems to arise naturally from the need to replicate a message. Nearly nine months after via was invented, we see the the very first usage of RT:

> RT @BreakingNewsOn: "LV Fire Department: No major injuries and the fire on the Monte Carlo west wing contained east wing nearly contained."

The tweet had exactly 140 characters, the length limit set by Twitter. This finding suggests that the invention of RT was a result of a previous variation being adapted to the constraints of the social environment.

The recycle icon, on the other hand, seems to have been created for the purpose of improving the existing variations, as it appeared explicitly in the discussion of which variation should be used on Twitter. The recycle icon variation was in fact the last one to be invented, in 2008 September. More details about the emergence process are described in [29]

4. THE RETWEETING CONVENTION

The seven variations, although started at various points in time with different contexts, were invented to act as a convention indicating retweeting and once introduced in Twitter were subsequently adopted by other users who were exposed to them. In this work, we ask *are there any features in the Twitter data set that would allow one to predict which variation would have been adopted by whom and when?* As we demonstrate shortly, the final reach of these variations does not seem to be strongly related either to the amount of time each variation had to grow or the rate at which it grew. Therefore, we sought to investigate the user-level traits and the social network characteristics that determine the adoption process.

The longitudinal tweet data set allows us to track how each of the variations was adopted. Figure 1 shows the time series of week-to-week usages over the first 3.5 years of Twitter's existence. Via started early with a slow growth pattern relative to the other variations, but ended with the second highest number of adopters. In contrast, RT started late but soon became a leading variation. The recycle icon, HT, and R/T continued to add new users, but their popularity seems to have nearly stabilized. Retweet and Retweeting once grew as fast or faster than any other variations, but never approached the reach of RT. In fact, they began losing popularity in late 2008, as the number of usages declined.

With respect to predicting the future popularity of a social convention, we focus on one specific problem, which we call the **Convention Prediction Problem** and define as follows: Suppose we are given a social network with records of users and times of adoptions, but information about which variation was adopted by user u at time t is hidden. How reliably can we infer which variation v the user chose to adopt?

The prediction problem takes into account a wide range of features including personal preference (e.g., in-degree and geo-location of adopters) as well as social correlation due to homophily and influence (e.g., the frequency and degree of exposures from friends). We chose these features because they correspond to potential causes identified in the research on adopting existing social norms [17] and the diffusion of other social phenomena [4]. In the remainder of this section, we consider the effect of some of the key features in the emergence of the retweeting convention. Then in Section 5, we describe our solution to the prediction problem, considering the combined effect of all features.

4.1 Personal Features

One way to distinguish the small group of users who adopted one of the variations of the retweeting convention from the typical Twitter user is a simple count of one's followers and followings. Adopters had higher in-degree than the average Twitter user. In fact, as can be seen in Figure 2, adopters have an order of magnitude more followers than the typical user. The figure also shows that RT and via were adopted by less popular users than the other variations. We hence consider *in-degree and out-degree of adopters* as one of the predictive features.

An in-depth examination revealed that in-degree distributions of the first 500 adopters were not particularly different from one variation to another [29]. This finding may suggest that when a variation becomes extremely popular, it starts to be adopted by users beyond the circle of highly-connected core users and reach more peripheral users—a finding that is in tune with the famous "diffusion of innovations" theory [37]. In contrast, adopters of the less popular variations may have never broken out of the core group of Twitter users who pay attention to the new trends and technologies of the service.

We also examined whether a user's geographical location affects her decision to adopt a variation. It is possible that the conventions differed from location to location because of language or cultural differences. In order to infer the geo-location of adopters, we em-

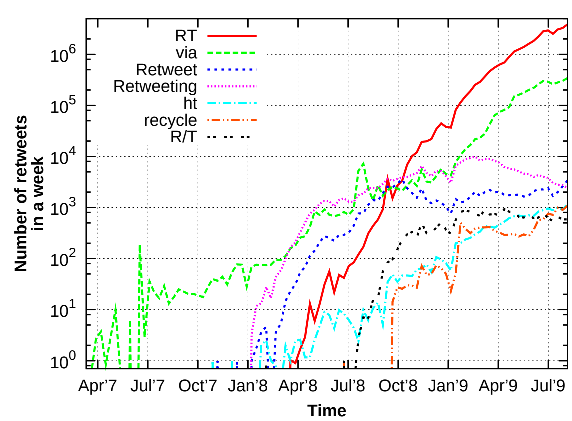


Figure 1: Number of usages of different variations over time

Country	Total adopters	RT	via	Retweeting	Others
US	1,239,708	0.668	0.285	0.021	0.025
UK	137,848	0.664	0.295	0.022	0.019
Brazil	112,465	0.746	0.234	0.014	0.006
Canada	89,803	0.656	0.298	0.023	0.023
Germany	48,575	0.632	0.316	0.033	0.019
Australia	46,438	0.659	0.294	0.023	0.023
Japan	28,859	0.731	0.234	0.005	0.030
Iran	27,330	0.622	0.315	0.024	0.038
France	18,934	0.600	0.341	0.042	0.018
India	17,449	0.623	0.320	0.033	0.024

Table 2: The fraction of different variations in top 10 countries

ployed two fields from user profiles: the location field and the time zone field. The location field is a free-text string entered by users, while the time zone field is a drop-down menu in Twitter. To resolve the free-text string in the location field, we used both the Yahoo! Map and the Bing Map APIs (Application Programming Interfaces) and obtained the country-level information for each user. The time zone field has a location name appended with the Unix Time offset, with which we can resolve the country individual users belong to. We annotated users with the geo-location information only when at least two of the three inference results matched. In this manner, we were able to retrieve the geo-location for 75.5% of all adopters.

Investigating the distribution of variations across geography, we found that geo-location had little impact in the spread of popular variations like RT and via. However, geo-location showed some variability for less popular variations like HT and recycle icon. Table 2 shows the top 10 countries based on the total number of adopters and the distribution of variations within that country. In all ten countries (and in most other countries not listed in the table), RT was the leading variation, followed by via. For the remaining variations, however, the ordering was not consistent. For instance, users in France and Germany preferred the recycle icon to Retweet, while users in Iran and Australia preferred HT much more than in other countries. In order to capture the effect of geo-location, we consider *the country of the adopters* as a predictive feature.

4.2 Social Features

One of the key factors that is essential in convention adoption is *exposures*. These are so crucial because of informational influence [20]; that is, users are made aware of the variation by their peers and as a consequence adopt the variation. This is especially true when conventions are first being established, because there is a utility simply in coordinating on the same variation [33]. Of course, it is likely that some users independently invented the same variation or learned about the variations through some external source. However, because retweeting is specific to Twitter, there is a high chance that adopters of this convention were exposed to the variation through their social contacts in Twitter. We call these users *internal adopters*.

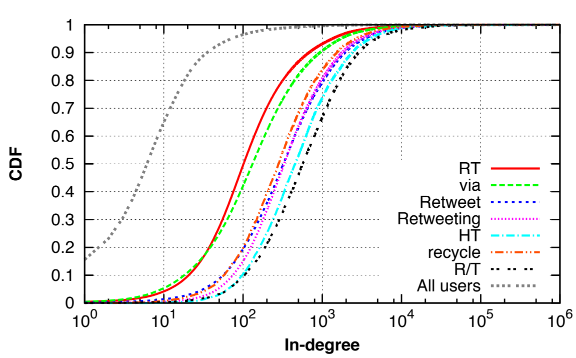


Figure 2: Comparison of in-degree of users

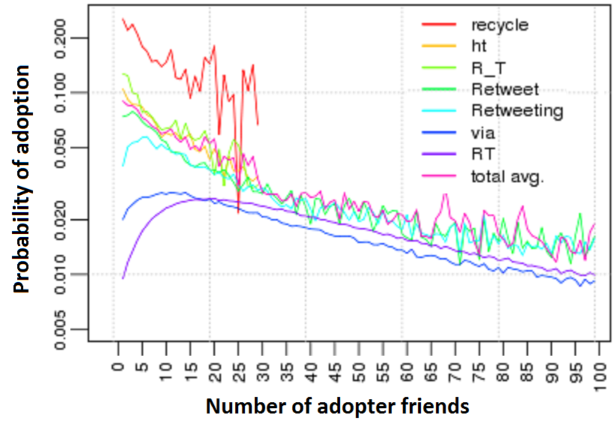


Figure 3: Exposure curve for seven variations

Table 3 shows the percentage of internal adopters. It should be noted that these percentages are likely an overestimate, as the graph of social contacts is based on a snapshot at the end of the data set in mid 2009, and connections in online social networks are much more likely to be added than deleted. Nonetheless, the proportion of adopters whom we assume learned about the convention internally is so high (96.71%) that it is unlikely the amount we overestimate the internal adoptions would be so substantial as to affect our conclusions. The three most popular variations (i.e., RT, via, and Retweeting) have the highest fraction of internal adopters. The overall high rate of internal adopters indicate that exposure is a key factor in the spread of a convention. From this observation, we later consider *frequency of exposure to a particular variation* and *the number of adopter friends* as prediction features.

The effect of exposure sometimes became stronger with repetition. Figure 3 shows a plot of the fraction of users who adopted a given variation after k of her friends adopted the same variation, as a function of k. RT and via show a two-phase pattern: a ramp-up to a peak followed by a decline, where the peak values are at k=20 and k=11, respectively. This observation indicates that the more a user is exposed to the variation up to these points, the more likely she will adopt it. The three lesser popular variations (i.e., HT, R/T, and recycling icon) do not have a ramp-up period and their peaks are at k=1. This means that for these variations repeated exposures were less effective than the first exposure.

Convention	Internal adopters
RT	97.93%
via	94.36%
Retweeting	95.56%
Retweet	89.17%
HT	89.32%
R/T	88.98%
recycle icon	81.97%
Total	96.71%

Table 3: Percentage of internal adopters for each variation

Convention	Switch-out	Switch-in
RT	23.30%	18.66%
via	18.13%	13.85%
Retweeting	73.61%	0.84%
Retweet	67.61%	0.67%
HT	30.94%	0.28%
R/T	57.81%	0.15%
recycle icon	35.67%	0.11%

Table 4: Loyalty of the convention adopters

Despite the fact that social exposure significantly increases the chance of adoption, it does not necessarily mean that the prediction problem is easy. The reason is two-fold. First, most adopters are exposed to not just one but multiple variations before adopting any of them. In fact, a mere 2.49% of the adopters were exposed to exactly one variation. Second, people do not always choose to use the variation that they are most frequently exposed to. Of those who adopted RT, 80.40% had been exposed to RT the most frequently. However, for the other variations, only 6.17%–30.39% of adopters had been exposed to that variation more than any other variation.

Amazingly, despite the fact that nearly every adopter was exposed to more than one variation, a great majority of the users (72.68%) only adopted a single variation. Most others (24.47%) switched from one variation to another once or used two different variations back and forth. Only a small fraction of users (2.85%) adopted three or more variations. In order to understand why users adopt multiple variations, we analyzed how frequently users switch between different variations.

We approach this by treating adoption of convention as a Markov process, in which we model the usage of each variation as individual states and switching from one variation to another as a transition between states. These individual decisions to switch from one variation to another on aggregate also reflect the proportion of the population moving from one variation to another. Let the transition probability from state i to state j be P_{ij} and let V denote the set of all variations. Then, we may calculate the probability of *switching-in* and *switching-out* of a particular variation i as $\sum_{j \in V} P_{ji}$ and $\sum_{j \in V} P_{ij}$, respectively. The former represents the attractiveness of the variation, while the latter represents the loyalty of an adopter upon employing the variation.

Table 4 shows the two transition probabilities. RT and via have relatively low switch-out probabilities; once a user employs these variations, he will likely continue to use it further. Retweeting and Retweet have high switch-out probabilities; once a user employs these variations, he is likely to switch to another variation in the future. This is potentially because Retweeting and Retweet are attractive at first because of their straightforward meaning, but are costly for subsequent usages because of their long length, given the 140 character limit. In terms of the switch-in probability, HT and recycle icon have the highest attractiveness upon exposure.

4.3 Global Features

We can also examine the switch-in and switch-out probabilities at different points in time. Figure 4 shows how the switch-out probability changes for each variation when we divide the data set into four different time windows. During the first phase until October 2008—a year and a half after the first usage of the retweeting convention—there is not much difference in the switch-out probabilities across the seven variations. Until this time, the concept of retweeting was still settling within the Twitter community and there was no agreement as to which variation should be used.

However, by the end of the second phase in February 2009, RT became much more popular than the other variations, with hundreds of thousands of uses compared to only thousands for the other variations. After this time period, the switch-out probability for redundant variations of RT such as Retweeting, Retweet, and R/T show a sharp increase. This trend reflects the collective action that users who previously adopted these variations quickly switched to a new variation (i.e., RT). Other variations also exhibit changes in the switch-out probability over time, albeit at a marginal level.

This collective switch out action can be observed clearly in the state diagram of the Markov model in Figure 5, which shows the normalized switch out probability from each state so that the sum of probabilities outgoing from each state adds up to one. A transition in the model indicates what fraction of adopters of a given variation ever switched out to another variation. A self-loop represents the fraction of adopters who used exactly one variation. Self-loops are more evident for popular variations than the less popular ones. When switch out happens, we find that RT and via are the most popular destinations. In particular, adopters of HT and recycle icon preferred to switch out to via, while adopters of all other variations preferred RT over via.

The time-varying switch-out probability implies that the choice of which variation to adopt is no longer a local decision but a global one, where individuals are affected by the phase of the entire convention spreading process. What seems to play a role here is the global popularity of each variation. In order to incorporate this finding, we later consider *the date of adoption* as a feature of the convention prediction problem.

5. CONVENTION PREDICTION PROBLEM IN TWITTER

In this section, we use features about a user's characteristics and behavior on Twitter (*personal*), the behavior of those they follow on Twitter (*social*), and the date of adoption (*global*) to predict the behavior of individuals who adopted the retweeting convention in Twitter. By examining how well we can predict an individual's decision to use one variation over another, we aim to understand which features play a key role in the adoption.

5.1 Features and Classifiers

The quality of any prediction algorithm depends on the choice of features. Based on the observations from the previous section, we considered the following features from the three categories described above: (1) *personal* features including the join date, geolocation, in-degree, out-degree, the number of tweets, and the number of URLs; (2) *social* features including the number of exposures to each variation and the number of friends who exposed the user to each variation; and (3) *global* features like the date of adoption. We

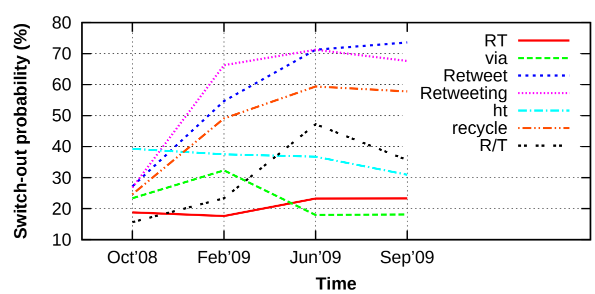


Figure 4: Switch out probability over time

Sample data entry	625, Germany, 321, 167, 252, 57, 54, 22, 5, 12, 1, 0, 4, 30, 18, 4, 6, 1, 0, 4, 665, RT

Table 5: An example of the data sample given to the classifier

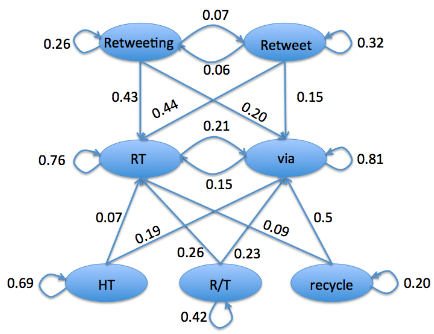


Figure 5: Transition probabilities between different variations (Probabilities below 0.05 are not drawn for visual clarity.)

consider the date of adoption because it correlates with the global popularity of the adopted variation.

Table 5 shows a sample data point. This sample represents a single adoption instance of an individual with different features at the time of adoption separated by comma in the order we introduced the features in the previous paragraph. The first number (i.e., 625) shows the join date of the adopter, in terms of number of days after Twitter was started. The next number indicates the location of the adopter, Germany in this case. The third and fourth numbers are the in-degree (321) and out-degree of the adopter (167), respectively. The numbers of posted tweets (252) and URLs (57) are the next two fields. The next seven fields are the total number of exposures to the seven different variations, which means the adopter was exposed 54 times to RT, 22 times to via, 5 times to Retweeting, and so on. Similarly, the next seven numbers show the number of adopter friends: 30 RT-adopter friends, 18 via-adopter friends, and so on. The last number represents the date of adoption in terms of number of days after Twitter was launched, 665 in this case, which is 40 days after the user joined Twitter. The final field shows the target variation of adoption (RT) that we want to predict.

In order to determine the discriminative power of each one of the 21 features, we computed the chi-square (χ^2) value and the information gain. Table 6 shows the order of all features based on the χ^2 value, where a larger value indicates a higher discriminative power. The order of the features were very similar when ranked by information gain (i.e., the Kullback-Leibler divergence), and in both cases the top three features span all three categories: global (date of adoption), social (the number of exposures to RT), and personal (the number of posted URLs).

Analysis of the discriminative power of the 21 features brings two key insights. First, the fact that the global feature (i.e., date of adoption) had the most predictive power suggests that the decision at the microscopic level—individuals deciding which variation to adopt—has more to do with the global popularity of the variations than its local popularity or a user's personal traits. Second, for each of the seven variations, the number of exposures had more predictive power than the number of adopter friends of that variation. This difference is important because most existing models of diffusion and adoption, such as the independent cascade model or the linear threshold model, are based on the distinct number of adopting friends rather than the sheer volume of exposures.

Rank	Feature	Type	χ^2 value
1	date of adoption	Global	300,666
2	# of exposures to RT	Social	106,627
3	# of posted URLs	Personal	80,728
4	# of exposures to via	Social	64,160
5	join date of the adopter	Personal	48,071
6	# of posted tweets	Personal	44,523
7	# of RT-adopter friends	Social	44,079
8	# of exposures to Retweeting	Social	43,100
9	# of exposures to HT	Social	42,807
10	# of exposures to Retweet	Social	36,604
11	# of exposures to recycle	Social	32,889
12	in-degree of the adopter	Personal	30,762
13	# of HT-adopter friends	Social	24,338
14	out-degree of the adopter	Personal	24,338
15	# of exposures to R/T	Social	23,370
16	# of Retweeting-adopter friends	Social	8,141
17	# of Retweet-adopter friends	Social	7,507
18	country of the adopter	Personal	4,476
19	# of via-adopter friends	Social	4,429
20	# of R/T-adopter friends	Social	203
21	# of recycle-adopter friends	Social	67

Table 6: Predictive features and their discriminative power

5.2 Classification Results

To test the combined effect of the 21 features described in Table 6, we built a classifier that predicts which variation a user adopts given the time of adoption. Predicting which of multiple variations a user will adopt (i.e., multi-class classification) is hard in this case because the classes are unbalanced. In total, 68.2% of all adoptions were RT, which means that a classifier that always predicts that a user will adopt RT will achieve at least that level of accuracy. In contrast, all other variations had much fewer adopters, for instance, recycle icon was only adopted by 0.2% of users. This means a classifier can achieve an accuracy of 99.8% for recycle icon by always predicting that the user did *not* adopt recycle icon, which is a very difficult baseline to improve upon.

For prediction, we used several classifiers in the WEKA machine learning toolkit [45] including SVM, Bayesian models, boosting, and decision trees. In particular, we used a popular ensemble classification technique, *bagging*, to achieve very high classification accuracy. Bagging aggregates the decision from a number of classifiers such that the accuracy is always at least as good as, and typically better than, the best classifier in the ensemble. In order to promote model variance, bagging trains each model in the ensemble using a randomly-drawn subset of the training set and then makes a prediction by taking an equally-weighted vote across all classifiers in the ensemble.

We tried many different combinations and subsets of the features and found that excluding 8 features (e.g., the number of adopter friends for each variation and the geo-location of the adopter) produced the best result. Table 7 shows the results of the multi-class prediction by using a random half of the data samples for training and the other half for testing. We correctly predicted 72.6% of the instances. Compared to the baseline of 70.4%, which is also shown in the table, this is a marginal improvement (<2.2%).

Variation	Baseline	Accuracy	Precision	Recall	AUC
RT	0.682	0.712	0.728	0.920	0.681
via	0.720	0.726	0.521	0.237	0.666
Retweeting	0.981	0.980	0.431	0.177	0.905
Retweet	0.986	0.985	0.343	0.084	0.801
HT	0.996	0.997	0.505	0.039	0.849
R/T	0.998	0.998	0.190	0.001	0.815
recycle icon	0.998	0.999	0.359	0.009	0.823
Weighted average	0.704	0.726	0.657	0.698	0.683

Table 7: Result of multi-class prediction for each variation

Variation	Accuracy	Precision	Recall	AUC
RT	0.613	0.607	0.631	0.662
via	0.607	0.606	0.601	0.650
Retweeting	0.591	0.589	0.618	0.641
Retweet	0.569	0.566	0.566	0.604
HT	0.823	0.828	0.815	0.896
R/T	0.773	0.770	0.772	0.838
recycle icon	0.815	0.831	0.802	0.898
Weighted average	0.610	0.607	0.615	0.657

Table 8: Result of predicting variations separately

As mentioned earlier, the classes are unbalanced and this makes it hard to see how much each feature helps to predict an adoption. To investigate whether the features were at all useful for predicting which variation a user would adopt, we selected all of the cases where a target variation and another variation were adopted on the same date. In this way, we ensured a balanced set of adoptions and controlled for the shifting popularity that we observed in the multi-class prediction problem. After selecting these cases, we performed the analysis with the same tool, method, and features (other than the date of adoption).

Table 8 shows the results of predicting variations separately. The classifier could achieve the accuracy of on average 22.0% improvement over the baseline, which is now 0.50 for all seven variations. The classifier could more accurately predict the adoptions for variations with fewer total adopters. In the case of HT, 82.3instances were correctly predicted, a 65This difference could be explained by the fact that adopters of these variations are in general more active and popular than the adopters of popular variations (Figure 2).

In order to understand the extent to which classification results are sensitive to any assumptions made, we repeated the prediction analysis with two stronger definitions of a social tie. Therefore, a user was not considered an adopter unless she had used a given variation at least three or five times. In addition, we considered the network structure based on user interactions rather than a simple follow link. In building the social network, we define user A to be linked to user B if and only if user A explicitly mentioned user B's name using "@username" in her tweets at least twice, following the definition of a "friend" in [8]. Directing a post to someone is called a *mention* in Twitter. Compared to a typical follow link, a "mention" link represents a stronger social relationship.

Figure 6 shows the prediction results under these different assumptions about a social link as well as the results for the standard definition of adoption and links as shown in Table 8. The accuracy when enforcing the 3-use threshold is better for the top five variations, but slightly worse for the bottom two variations. This is because having a stronger definition of adoption lead to fewer samples, and this meant the less popular variations had less data

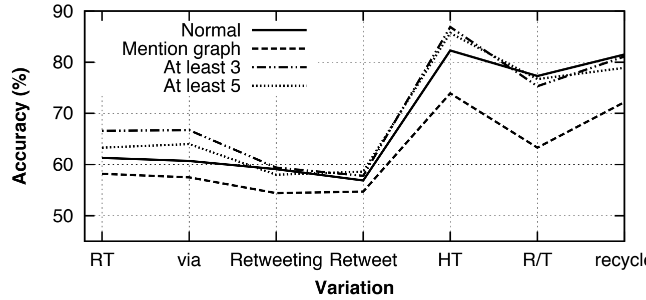


Figure 6: Prediction accuracy under different assumptions

for training. The overall prediction accuracy improved by an additional 5.4% compared to the previous prediction. Continuing to enforce a stronger definition (i.e., the 5-use threshold), however, did not continue to increase accuracy for the popular variations.

The prediction results for limiting the diffusion over the mention links as opposed to regular follow links are worse than the original prediction for all seven retweeting variations. This suggests that enforcing a stronger definition of social relationship does not necessarily improve the prediction power. This only resulted in cases where we could not explain the adoption process through social network. What seems more important is the exposure to the social convention itself, rather than the information source in the spread of retweeting convention in Twitter.

6. IMPLICATIONS

Our analysis suggests that the most important feature in predicting the user-level adoption was the global trend, as opposed to social or personal features. This is surprising, as one would expect that for this domain—where the convention is only useful within Twitter and is most likely to be learned about from one's peers on Twitter—the importance of local features would be greater.

There are, however, a number of possible reasons why the global trend would outweigh personal features. One could be that the Twitter culture was changing in ways over time that we did not measure or were unobservable. For instance, a different kinds of people could have joined Twitter, where the first group of people preferred via and later people preferred RT. Or it could be that independent of the new users, the culture within Twitter or even outside of Twitter was changing in some way that led users to prefer one variation over another (e.g., an increase in the sharing of breaking news and URLs). If this is true, it raises questions about how these changes in the culture are affecting the Twitter population *en masse* if it is not evident in the interactions between individuals.

We could also reasonably assume that some users prefer to adopt the more globally popular convention as opposed to the more locally popular one. Users can get the global view in different ways; for example popular users and celebrities who are followed by many users could represent the global view. Users can also be exposed to a set of random tweets shown in the front page of Twitter or from the search results.

It is worth pointing out that just because a global feature proved to be the strongest predictor, it does not mean that the most popular variation at the moment will necessarily receive strictly more adopters. Twitter was not a "winner-take-all" market. In fact, the dominant variant was no longer the most popular after a few months; the top ranked variation was initially via for nearly a year, then switched back and forth between Retweeting and via for half a year until RT became the dominant one in late 2008 (Figure 1). This means that there may be a process of constant re-evaluation within the Twitter community, through which the rank of the variations could change.

A second key results is that the number of exposures had more discriminative power than the number of adopter friends. This is interesting because other work has suggested that the number of distinct adopter friends is crucial for predicting one's adoption [28]. However, we saw that the number of adopter friends was less important than the volume of repeated exposures. With respect to the effect of repeated exposures in adoptions, Romero and colleagues found a similar positive effect in the spread of political hashtags in Twitter [38]. They describe this as "complex contagion" and explain that when a particular behavior is controversial or contentious, people may need more exposure to it from others before adopting it themselves. In the case of the retweeting convention, people may also need to be exposed to this new practice multiple times before they start to adopt it. This seems especially likely, since in the development of a social convention, as framed in the coordination problem [42], users require even more "social proof" before deciding to adopt the convention.

7. CONCLUSION

How do social conventions emerge and settle in human societies? In this paper, we presented one of the first studies using large-scale empirical data to study the emergence of a social convention. The convention we study is not based on simulations or experiments in a controlled environment as in most existing work, but is an actual practice that was adopted by more than 2 million people in the popular social network Twitter. The retweeting convention we studied is specific to Twitter, hence serves as a relatively clean case study in which we have records of nearly all usages and exposures of the convention.

Given the need to establish a convention for retweeting, a number of variations of the convention emerged organically in Twitter. As we saw in the tweets of the early adopters, users actively engaged in the discussion of which variation should be used. Throughout the course of the first few years, different variations competed to reign as a leading practice and in this process certain variations (e.g., longer ones) showed to be more easily abandoned than others. When it comes to individual's decision to adopt a variation, a number of factors such as social exposures and the date of adoption played an important role. However, predicting the adoption process at a user level turned out to be challenging.

There are several exciting directions to pursue. Since the features we considered only represent simple personal traits such as the node degree and geo-location, we believe more nuanced features about the adopters (e.g., socio-economic factors) would have the potential to improve the prediction. For instance, Pennacchiotti and Popescu demonstrated that users' political affiliation, ethnicity, and affinity for a particular business can be inferred in Twitter [35], and these features may be useful for determining which retweeting variation a user will adopt. Furthermore, it is possible that by leveraging the features we studied as well as other potential features, one may be able to extrapolate to the macroscopic problem of predicting which variation would come to be the most popular based on early usages.

Acknowledgements

Meeyoung Cha was supported by Basic Science Research Program through the National Research Foundation, Korea (2011-0012988).

8. REFERENCES

[1] I. Ajzen and M. Fishbein. The prediction of behavior from attitudinal and normative variables. *Journal of Experimental Social Psychology*, 6:466–487, 1970.

[2] S. E. Asch. Opinions and Social Pressure. *Scientific American*, 193:31–35, 1955.

[3] R. Axelrod. An Evolutionary Approach to Norms. *American Political Science Review*, 1986.

[4] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic. The Role of Social Networks in Information Diffusion. In *proc. of the International World Wide Web Conference*, 2012.

[5] F. Barth. *Process and Form in Social Life*. Routledge & Kegan Paul, 1981.

[6] R. Baumeister and M. Leary. The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation. *Psychological Bulletin*, 1995.

[7] B. Becker and G. Mark. Social conventions in collaborative virtual environments. In *proc. of the International Conference on Collaborative Virtual Environments*, 1998.

[8] D. M. R. Bernardo A. Huberman and F. Wu. Social networks that matter: Twitter under the microscope. *First Monday*, 14(1), Jan. 2009.

[9] Bibb Latané and S. Wolf. The social impact of majorities and minorities. *Psychological Review*, 88:438–453, 1981.

[10] d. boyd, S. Golder, and G. Lotan. Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In *proc. of the Hawaii International Conference on System Sciences*, 2010.

[11] R. Boyd and P. J. Richerson. Why does culture increase human adaptability? *Ethology and Sociobiology*, 16:125–143, 1995.

[12] B. P. Buunk and A. B. Bakker. Extradyadic sex: The role of descriptive and injunctive norms. *Journal of Sex Research*, 32(4):313–318, Jan. 1995.

[13] D. Centola. The Spread of Behavior in an Online Social Network Experiment. *Science*, 329(5996):1194–1197, Sept. 2010.

[14] M. Cha, H. Haddadi, F. Benevenuto, and K. Gummadi. Measuring User Influence in Twitter: The Million Follower Fallacy. In *proc. of the International AAAI Conference on Weblogs and Social Media*, 2010.

[15] N. Christakis and J. H. Fowler. The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine*, 2007.

[16] R. B. Cialdini. *Influence: Science and practice*. HarperCollins, 3 edition, 1993.

[17] R. B. Cialdini and M. R. Trost. Social Influence: Social norms, conformity, and compliance. In D. T. Gilbert, S. T. Fiske, and L. Gardner, editors, *The Handbook of Social Psychology*, pages 151–192. McGraw-Hill, New York, 1998.

[18] E. Cohen-Cole and J. M. Fletcher. Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of Health Economics*, 27:1382–1387, 2008.

[19] J. Delgado. Emergence of social conventions in complex networks. *Artificial Intelligence*, 141(1-2):171–185, 2002.

[20] M. Deutsch and H. B. Gerard. A study of normative and informative social influences upon individual judgment. *Journal of Abnormal Psychology*, 51:629–636, 1955.

[21] D. C. Feldman. The development and enforcement of group norms. *The Academy of Management Review*, 9(1):47–53, 1984.

[22] L. Festinger. A theory of social comparison processes. *Human Relations*, 7(2):117–140, 1954.

[23] J. H. Fowler and N. A. Christakis. Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences*, 107(12):5334–5338, Mar. 2010.

[24] N. E. Friedkin. Norm Formation in Social Influence Networks. *Social Networks*, 23:167–189, 2001.

[25] H. B. Gerard. The Anchorage of Opinions in Face-to-Face Groups. *Human Relations*, 7(3):313–325, Aug. 1954.

[26] H. B. Gerard, R. A. Wilhelmy, and E. S. Conolley. Conformity and group size. *Journal of Personality and Social Psychology*, 8:79–82, 1968.

[27] S. Goyal and M. C. W. Janssen. Non-Exclusive Conventions and Social Coordination. *Journal of Economic Theory*, 77(1):34–57, 1997.

[28] C. A. Insko, R. H. Smith, M. D. Alicke, J. Wade, and S. Taylor. Conformity and group size: The concern with being right and the concern with being liked. *Personality and Social Psychology Bulletin*, 11:41–50, 1985.

[29] F. Kooti, H. Yang, M. Cha, K. Gummadi, and W. A. Mason. The emergence of conventions in online social networks. In *proc. of the International AAAI Conference on Weblogs and Social Media*, 2012.

[30] J. Levine and R. L. Moreland. Progress in Small Group Research. *Annual Reviews in Psychology*, 41:585–634, 1990.

[31] S. Milgram, L. Bickman, and L. Berkowitz. Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology*, 13(2):79–82, 1969.

[32] T. M. Newcomb. *Personality and social change*. Dryden, 1943.

[33] K.-D. Opp. How do norms emerge? An outline of a theory. *Mind & Society*, 3(2):101–128, 2001.

[34] E. Ostrom. Collective Action and the Evolution of Social Norms. *The Journal of Economic Perspectives*, 14:137–158, 2000.

[35] M. Pennacchiotti and A.-M. Popescu. Democrats, Republicans and Starbucks Afficionados: User Classification in Twitter. In *proc. of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2011.

[36] A. Riley and P. J. Burke. Identities and Self-Verification in the Small Group. *Social Psychology Quarterly*, 58(2):61–73, 1995.

[37] E. M. Rogers. *Diffusion of Innovations*. Free Press, 1983.

[38] D. M. Romero, B. Meeder, and J. Kleinberg. Differences in the Mechanics of Information Diffusion Across Topics: Idioms, Political Hashtags, and Complex Contagion on Twitter. In *proc. of the International World Wide Web Conference*, 2011.

[39] M. Salganik, P. S. Dodds, and D. J. Watts. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, 311:854–856, 2006.

[40] B. Savarimuthu, S. Cranefield, and M. Purvis. Role Model Based Mechanism for Norm Emergence in Artificial Agent Societies. In *proc. of the AAMAS '07 Workshop on Coordination, Organization, Institutions and Norms in Agent Systems*, 2007.

[41] M. Sherif. *The psychology of social norms*. Harper, 1936.

[42] Y. Shoham and M. Tennenholtz. On the emergence of social conventions: modeling, analysis, and simulations. *Artificial Intelligence*, 94(1-2):139–166, 1997.

[43] D. M. Taylor and W. Louis. Terrorism and the quest for identity. In F. M. Moghaddam and A. J. Marsella, editors, *Understanding terrorism: Psychosocial roots, consequences, and interventions*, volume 343, pages 169–185. American Psychological Association, 2004.

[44] A. Walker and M. Wooldridge. Understanding the emergence of conventions in multi-agent systems. In *proc. of the International Conference on Multi-Agent Systems*, 1995.

[45] WEKA website. http://www.cs.waikato.ac.nz/ml/weka.

[46] D. Wilkes-Gibbs and H. Clark. Coordinating beliefs in conversation. *Journal of Memory and Language*, 31(2):183–194, 1992.

[47] M. P. Zanna and J. M. Olson. Individual differences in attitudinal relations. In M. P. Zanna, E. T. Higgins, and C. P. Herman, editors, *Consistency in social behavior: The Ontario Symposium*, volume 2, pages 75–103. Erlbaum, 1982.