Information Processing and Management 56 (2019) 102070

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

The reflection of offline activities on users' online social behavior: An observational study

Seyed Amin Mirlohi Falavarjani[a], Fattane Zarrinkalam[b], Jelena Jovanovic[c], Ebrahim Bagheri[b,*], Ali A. Ghorbani[a]

[a] Faculty of Computer Science, University of New Brunswick, Fredericton, NB, Canada
[b] Department of Electrical, Computer and Biomedical Engineering, Ryerson University, Toronto, ON, Canada
[c] Department of Software Engineering, University of Belgrade, Belgrade, Serbia

ABSTRACT

The ever increasing presence of online social networks in users' daily lives has led to the interplay between users' online and offline activities. There have already been several works that have studied the impact of users' online activities on their offline behavior, e.g., the impact of interaction with friends on an exercise social network on the number of daily steps. In this paper, we consider the inverse to what has already been studied and report on our extensive study that explores the potential causal effects of users' offline activities on their online social behavior. The o*bjective of our work* is to understand whether the activities that users are involved with in their real daily life, which place them within or away from social situations, have any direct causal impact on their behavior in online social networks. Our work is motivated by the theory of *normative social influence*, which argues that individuals may show behaviors or express opinions that conform to those of the community for the sake of being accepted or from fear of rejection or isolation. We have collected data from two online social networks, namely Twitter and Foursquare, and systematically aligned user content on both social networks. On this basis, we have performed a *natural experiment* that took the form of an *interrupted time series with a comparison group design* to study whether users' socially situated offline activities exhibited through their Foursquare check-ins impact their online behavior captured through the content they share on Twitter. Our main findings can be summarised as follows (1) a change in users' offline behavior that affects the level of users' exposure to social situations, e.g., starting to go to the gym or discontinuing frequenting bars, can have a causal impact on users' online topical interests and sentiment; and (2) the causal relations between users' socially situated offline activities and their online social behavior can be used to build effective predictive models of users' online topical interests and sentiments.

1. Introduction

The understanding of human behavioral patterns at both macro (Khatua, Khatua & Cambria, 2019; Kropivnitskaya, Tiampo, Qin & Bauer, 2017; Paul, Dredze & Broniatowski, 2014) and micro (Mueller, Jay, Harper & Todd, 2017; White & Horvitz, 2017; Zarrinkalam, Kahani & Bagheri, 2018) levels has become a major topic of exploration due to the growing availability of human activity traces. The collected trace data are not limited to users' geographical location, but also capture users' sentiments, opinions, and communication. For instance, while fitness trackers, navigation systems, and cell phone location sharing services have long had access to the geographical position of the users, messaging and social content sharing platforms are gaining ever increasing access to user generated content that can shed light on users' behavioral patterns from a multitude of aspects. Recent developments have explored the possibility of extracting actionable insight from user activity traces for different application areas such as advertising (Li, Lu, Mei, Wang & Pandey, 2015), healthcare (Greaves, Ramirez-Cano, Millett, Darzi & Donaldson, 2013), and customer care (Sulistya, Sharma & Lo, 2016), just to name a few. These works often view user activity traces as a source for supervised and unsupervised extraction of features for user behavior modeling and prediction (Hofman, Sharma & Watts, 2017; Zhao et al., 2011),

* Corresponding author.
E-mail address: bagheri@ryerson.ca (E. Bagheri).

https://doi.org/10.1016/j.ipm.2019.102070
Received 15 November 2018; Received in revised form 23 April 2019; Accepted 1 July 2019
Available online 15 July 2019
0306-4573/ © 2019 Elsevier Ltd. All rights reserved.

which could then be employed for systematic decision making and recommendation.

More recent works have taken a step beyond finding useful correlations for building predictive models by focusing on causal relations between various phenomena and user behavior. The lack of control over the distribution of subjects into comparable groups when dealing with historical online trace data has prevented researchers from using Randomized Controlled Trials (RCTs) and motivated them to explore alternative methods in the form of *observational studies* as well as *natural and quasi-experiment designs*. Interesting findings have already been reported based on such study designs that include user dietary behavior (Mejova, Haddadi, Noulas & Weber, 2015), psychological states (Althoff, Clark & Leskovec, 2016), and suicide ideation (Choudhury, Kiciman, Dredze, Coppersmith & Kumar, 2016). These forms of causal studies have also explored the impact of users' online behavior on their offline activities. For instance, Stück, Hallgrímsson, Steeg, Epasto and Foschini (2017) reported that Fitbit users who were active on the Fitbit social network and suffered from diabetes were six-times more likely to increase their daily steps for each additional social connection they had. Similarly, Althoff, Jindal and Leskovec (2017) found that each addition of a new social connection on the Argus app increased the number of steps a user took per day by 400 steps (7% increase). Such studies establish that, under some conditions, online presence and online behavior can shape a person's offline activities.

1.1. Study objective

The *objective of our work* in this paper is to understand whether users' offline activities, specifically those that place the user within or remove the user from a social situation, can influence their online behavior.[1] In order to capture both online and offline activities of users, we have curated a large dataset that aligns users' activities on Twitter and Foursquare. We collected users' online behavior in the form of their published tweets to analyze their topical interests and sentiments. Furthermore, we collected Foursquare checkins as they indicate users' offline activities based on the locations they have visited. We systematically study, through a *natural experiment*, whether embarking on or abandoning an offline activity such as frequenting a bar or quitting a gym, has any causal impact on users' online topical interests and sentiment states.

For example, in order to show the impact of frequenting a bar on online topical interests of a sample twitter user, who we name @ janedoe for anonymity, in Fig. 1, we show the distribution of her top-10 topical interests in four consecutive months, i.e. April, May, June, and July of 2016. Based on the check-in information of @janedoe in Foursquare, we know that she was not going to bars in April and May of 2016, but then an 'interruption' happened and she started going to bars in June and kept going to a bar every week for the next seven weeks. As such, we know that the user has actively embarked on some offline real-world activity (in this case starting to go to bars, which exposes the user to a *new social situation*), that she was not involved in before. We are now interested in examining if this new offline activity and the social exposure associated with it had any impact on the users' online topical interests. Let us assume that there were 300 topics actively discussed on Twitter in the examined time period (April–July 2016). The red dots in Fig. 1 show the top-10 topical interests of the user in each month (for details on determining users' topical interests see Section 4.3). The blue dots, in contrast, show the top-10 topics that were of interest to the general community in each month. As shown in Fig. 1, in the first two months, the topical interests of @janedoe were quite different from the topics of interest of the general community (blue and red dots do not align). However, after the 'interruption' happened, that is, the user started to go to bars in June, it can be observed that four of her topics of interests, i.e., z_8, z_{97}, z_{134} and z_{100} became very well aligned with the topics that the general community were interested in. Furthermore, in July, while the user was still going to bars, the alignment of her topical interests with the community became even more prominent (7 out of 10 of her top interests became aligned with the general community).

The objective of our work is to systematically study this observed pattern to see if it can be generalized across a larger group of users and for different types of offline activities with the ultimate goal of identifying causal relations between offline activities and online behavior.

1.2. Contributions

Understanding causal relations between users' offline activities and their online behavior, that is, how offline activities may affect online behavior, has practical applications in marketing, product recommendation, and customer retention, among others. The literature has already shown that customer satisfaction and quality of user interaction transcend beyond quality of the services and products and is also dependent on factors such as mood (Gardner, 1985), affective responses (Bagozzi, Gopinath & Nyer, 1999), persuasiveness (Petty, Cacioppo, Sedikides & Strathman, 1988), and cognitive and topical interpretation of needs (Petty, Priester & Wegener, 1994). Therefore, in our study, we explore whether causal relations can be observed between users' offline activities that expose the user to social situations and their online expression of topics of interest and sentiments. The contributions of this work can be enumerated as:

1 We show how the alignment of content from two social networks, namely Twitter and Foursquare, can provide the means to study users' offline and online behavior in tandem in the context of *observational studies*;
2 We introduce and motivate two metrics to model users' online behavior, namely (i) *social alignment*: the degree to which a users' online behavior is inline with the general public both in terms of topics and sentiments, and (ii) *social convergence*: the extent of

[1] When referring to user's online behavior, we are specifically interested in user's topical interests and sentiment states over time. We acknowledge that the term 'online behavior' can have broader implications but we limit its scope to users' topical interests and sentiments in this paper.

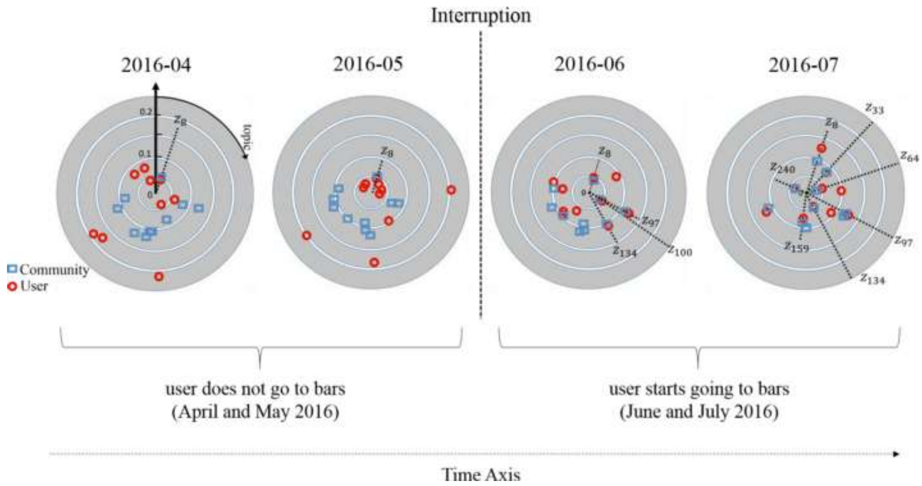

Fig. 1. Topical interests of user @janedoe and their alignment with the topics of interest of the overall community. Topics are placed around the radar and denoted by z_i, whereas contribution to each topic is specified by distance from the center. Topic alignments between the user and the community are marked with dotted lines.

change in users' online behavior towards or away from the community's overall behavior over an examined time period;

3 We employ the introduced metrics to study the *causal* impact of users' offline activities, specifically those that place the user within or remove the user from a social situation, on their online behavior as it relates to their topical interests and sentiments towards active topics on Twitter;
4 We show that it is possible to train machine learning-based predictive models solely based on users' offline activities in order to accurately predict the users' online topical interests and sentiments.

Briefly, we found that embarking on any of the examined offline activities can lead to change in users' online topical interests and can have significant influence on their topical sentiment. Based on a dataset of 48M Twitter posts and 6M Foursquare check-ins from 17,220 users, our findings show that users who actively embarked on an offline activity that was socially situated significantly changed their online behavior compared to two control groups: (1) a matched control group who did not take up the same offline activity and (2) the same users present in the treated group but in a different time interval. By studying these changes, we have found that those users who embarked on an offline activity, posted more tweets about topics that were of interest to the general public and their sentiment towards these topics was quite aligned with the sentiments shared by the community.

2. Related works

Social media and online social networks have been recognized as a large-scale 'sensor' that can provide insights into people's activity, behaviors, thoughts, emotions, and health. As a result, they have already been used as a viable source of information for observational studies and assessing behavioral characteristics of users.

Several observational studies have focused on social media and social networks by performing linguistic analysis of the user generated textual content to discover causal relations in the observed data (Oktay, Taylor & Jensen, 2010). For example, in the context of health-related data, Choudhury and Kiciman (2017) and Choudhury et al. (2016) have utilized mental health related content of semi-anonymous support communities on Reddit as a data source to perform an observational study with a statistical matching methodology in order to understand the transition from mental illness to suicide ideation and measure how the language of comments in Reddit mental health communities influence the risk of suicide ideation in the future. In Ernala et al. (2018), the objective of the researchers was to study the effect of disclosing a mental illness on the user's followers. They used Twitter data and found that, after a user disclosed her mental illness, the behavior of the user and the followers showed a significant correlation with respect to Schizophrenia symptoms. Further, Saha, Weber and Choudhury (2018) have examined the importance of counseling for students who have experienced death incidents while in college by evaluating student comments from Reddit communities. By comparing two groups of students, one consisting of students who had been exposed to counseling (commented on counseling recommendations) and the other of those who had not, Saha et al. found positive psychological effects of counseling.

Choudhury, Sharma and Kiciman (2016) have investigated how food related content on Instagram contributed to food choices and dietary patterns in "food deserts", that is, areas with poor access to healthy food. They found that: (1) content shared by users in food deserts, indicated consumption of food that were higher in fat, cholesterol, and sugar compared to non-food desert areas; (2) food desert posts were less frequently tagged with fruit and vegetable names, and (3) there were not only varied nutritional differences between food deserts and non-food deserts in different regions, but even the associated ingestion language showed variations. Ebrahimi, Phan, Dou, Piniewski and Kil (2016) have studied a health social network that tracked physical activities, biomarkers, and the users' posts to investigate how users' posts could influence physical activity behavior at the network level. They validated their findings by looking at users' actual medical progress and documented levels of exercise. While most of the work in the literature analyzed health content solely based on the linguistic analysis of user generated textual content, Althoff et al. (2017) have studied the causal effect of social connections (instead of user generated content) in a health application on the increase of users' online in-application activity and users' offline real-world physical activity. They have shown that: (1) the creation of new social connections increased user online in-application activity, user retention, and user offline real-world physical activity; (2) social influence accounted for 55% of the observed changes in user behavior; and (3) individual edge formations in the social network led to significant increases in physical activity.

Due to the popularity of Twitter, many researchers have focused on utilizing Twitter as a low-cost, large-scale, and rich data source for the purpose of observational studies (Paul & Dredze, 2011; Reis & Culotta, 2015). For example, Choudhury, Gamon, Counts and Horvitz (2013) have proposed a variety of measures such as language, emotion, style, ego network, and user engagement on Twitter to characterize depressive behavior and consequently predict likelihood of depression of individuals. Similarly, in Choudhury, Counts and Horvitz (2013), the authors have investigated the influence of childbirth on behavior and mood indicators of mothers on Twitter and have constructed a statistical model to predict significant postnatal behavioral changes. Posts and network data of Twitter users have also been analyzed by Murnane and Counts (2014) to reveal the distinguishing factors between those who failed in their smoking cessation and those who did not. Reis and Culotta (2015) have inferred causality between exercise and mental health (Anger, Depression, Anxiety) for Twitter users using observational studies with statistical matching methods. They found that those who exercised regularly had significantly fewer tweets expressing depression or anxiety while there was no significant difference in rates of tweets expressing anger. Recently, Olteanu, Varol and Kiciman (2017) have performed a quantitative analysis of Twitter data to answer some open questions related to the relationship between the experiences that users have already shared on Twitter, and the experiences they were likely to mention in the future. Celli, Ghosh, Alam and Riccardi (2016), proposed a new method for finding the relationship between personality types and communication styles of Twitter users and the types of contents they share on Twitter. They found that *depressed* communication styles were correlated with sharing of indignation articles. Based on this finding, the authors built a predictive model to predict the users' mood in the future.

Our work presented in this paper distinguishes itself from the literature in two main ways: (1) the work in the literature that compared users' online and offline behavior often relied on users' social content to model their online behavior, whereas the users' offline activities were collected through other means such as health monitors or fitness trackers. The work in this paper is among the few to propose a systematic way for using two separate social networks (Twitter and Foursquare) to collect users' online and offline activities; and, (2) while there have been several works that studied the impact of users' online behavior on their offline activities, there are few, if any, that have looked at how users' offline activities impacted their online behavior. Our work is among the first to use an observational study, in the form of a natural experiment, to systematically explore causal effects of users' offline activities on their online behavior.

3. Research questions

The main objective of our work is to systematically study whether a change in users' offline activities that places them within or removes them from social situations have any quantifiable impact on their online topical interests and sentiment. To this end, we introduce two central concepts, namely (1) *social alignment*, which addresses a user's tendency to agree with the public topical interests and sentiments, and (2) *social convergence*, which portrays a user's likelihood to shift towards the public interests and sentiments. These form the basis for our two main research questions (RQs):

RQ1 Would embarking on or abandoning an offline activity, which leads to a user's exposure to or retrieval from social situations, impact the user's degree of social alignment with public interests and sentiments?

RQ2 Would embarking on or abandoning an offline activity, which leads to a user's exposure to or retrieval from social situations, cause the convergence of the user's interests and sentiments towards that of the general public?

Each of the two RQs is further broken down as shown in Table 1.

The underlying theoretical assumption of our work has been motivated by existing work in *social influence* (Cialdini Robert, 1993). Several researchers have already argued that it is possible for individuals to be influenced by the behavior, thoughts and opinions of those they are surrounded by. For instance, Cialdini, Reno and Kallgren (1990) observed that college students who had just observed other students littering are more likely to litter themselves compared to those students who have just seen other students picking up litter from the ground and placing it in the trash can. This finding is fully aligned with Bandura's theory of social learning, which posits that people tend to learn from one another, through observation, imitation, and modeling (Bandura, 1977). Social influence can happen for various reasons, including the need to avoid being rejected by others within a social context. This type of social influence that leads to social conformity is known as *normative social influence*. Morton and Gerard (1955) discuss that individuals are often subject to normative social influence when they desire to feel the sense of belonging or urge the need to become accepted as a

Table 1
Overview of research questions.

	Topical interests	Topical sentiments
Social alignment	RQ1.1 Social Topic Interests Alignment	RQ1.2 Social Topic Sentiment Alignment
Social convergence	RQ2.1 Social Topic Interests Convergence	RQ2.2 Social Topic Sentiment Convergence

part of a group. In the context of normative social influence, it is possible for the individual to adopt the influence of the group to avoid social rejection and isolation without an active intent of conforming. Accordingly, we are interested in exploring whether initiating offline activities that place the user within a social context, such as starting to go the gym, bars or restaurants, or terminating such activities (e.g., abandoning the gym, bars or restaurants) have any impact on the user's topical and sentimental expressions. Our hypothesis, based on the normative social influence literature, is that once users start to partake in offline activities that expose them to social situations, they are more likely to (intentionally or unintentionally) adjust their topical and sentimental expressions to meet the opinions of the community. On the other hand, those users who stop practicing offline activities that include social exposure are less likely to be impacted by social influence and hence would have the tendency of being more distant from the topical and sentimental expressions of the broad community.

Finally, based on the outcomes of our two research questions and the associated hypotheses, we are also interested in examining whether users' offline behavior can serve as discriminative features to *predict* users' online topical interests and sentiments. This is primarily because if a causal impact exists from the users' offline activities onto their online behavior, then it should be possible to use the users' offline activities to accurately predict their online behavior. We will report on our experience in building such a predictive model and its predictive power in determining users' online behavior solely based on the users' observed offline activities.

4. Research methodology

The research methodology used in our work is based on a *natural experiment* (DiNardo, 2008). Specifically, in our case, the natural experiment takes the form of an interrupted time series design with a treated and a control group. Given the goal is to discover the effect of offline activities on users' online behavior, the *treatment* in our natural experiment is some offline activity and its *outcome* is users' online behavior. The *treated group* includes users who embark on or abandon an offline activity, which impacts the exposure of the users to or their retrieval from a social situation, after a certain period of time, whereas the *control group* is formed of users for whom such engagement with an offline activity was not observed. To sum up, our experiments consist of four major elements namely, (1) the dataset used for conducting the natural experiment, (2) the treatment and the treated group, (3) the control groups, and (4) the outcome, each of which will be explained in detail in the following.

4.1. Description of the dataset

Considering the objective of our study, our dataset needed to include data on users' offline activities and their online behavior. We collected data from Twitter to represent users' online behavior and Foursquare to capture their offline activities. In order to align the content from the two social networks and link users across the two platforms, we utilized the Swarm[2] application. This application, which was launched on May 2014 by Foursquare, allows users to share on Twitter and Facebook the places they visit based on Foursquare venues. This made it possible to identify users who were active on both Twitter and Foursquare and link their offline and online activities.

To create a dataset for our study, we gathered data from October 2014 to April 2017 using the Twitter API. In order to identify users who had Foursquare check-ins, we processed the collected tweets to identify those with relevant Swarm information. Having identified those users who were on Twitter and were also sharing their location using Swarm, we only retained those users who had Swarm check-ins in at least 10% and at most 50% of their tweets to ensure (1) we had enough check-in information for each user and (2) there were other content posted by the user on Twitter beside check-ins. We also eliminated users who tweeted in languages other than English. This resulted in 17,220 users who were both present on Twitter and also frequently posted their location information using Swarm.

After downloading tweets using Twitter API, we extracted venue IDs from tweets related to the Swarm application. The venues where the users had checked in were then pooled using Foursquare API and categorized into four venue categories as explained in Table 2. Moreover, Fig. 2 shows the distribution of the four different check-in categories on a per user basis. Our observation was that the use of Swarm and Foursquare check-ins increased over time. Further our examination of the distribution of check-in counts over the examined time period showed that restaurant check-ins were the most popular and sport check-ins were the least frequent.

The fact that our dataset was collected over more than a three-year time period, excludes any bias towards certain topics or sporadic user behavior. In terms of size, our dataset, which consists of 17,220 users, 6523,257 check-ins at 211 Foursquare venue subcategories and 48,672,327 tweets, is comparable to the datasets used in similar studies, such as the dataset described in White and Horvitz (2017) which was collected from a commercially closed platform and the Twitter dataset used in Garimella, Morales, Gionis, and Mathioudakis (2017), among others. It should be noted that all personally identifiable information were anonymized during the data collection.

4.2. Treatment and treated group

The *treatment*, in our study, is defined as participation in or abandoning of an offline activity, which impacts the exposure of the users to or their retrieval from social situations. The *treated group* is the group of users who are impacted by the treatment. More precisely, we define the treated group as those users who can be identified as impacted by the treatment based on their Foursquare

[2] https://www.swarmapp.com/.

Table 2
Foursquare categories considered for each venue category. * is the wildcard symbol and (*) is a controlled wildcard where we selectively choose only sports related keywords from the list of categories.

Venue category	Count	Foursquare categories
Shopping	33	*Mall, *Store, *Shop, *Market, *Boutique
Bar	22	*Bar, *Pub
Restaurant	131	Restaurant, Buffet, Noodle House, Taco Place, Pizza Place, Steakhouse, BBQ Joint, Burger Joint
Sport	25	*Gym, (*)Studio, (*)Rink, (*)Center, Pool, (*)Club, (*)Court, (*)Field, (*)Ground

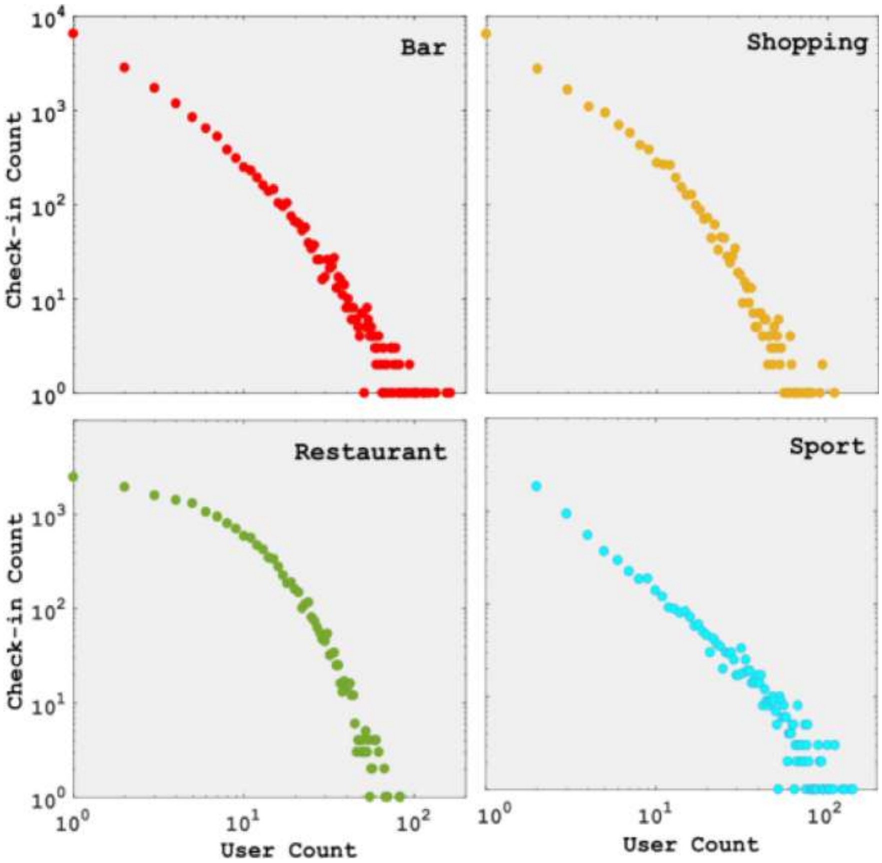

Fig. 2. Distribution of check-ins per user on a log-log scale.

checkins. We study two treatments, namely (1) embarking on and (2) abandoning an offline activity. Each of these two treatments are studied in the context of four venue categories, which include Bar, Restaurant, Shopping, and Sports. The specific Foursquare categories that were considered are shown in Table 2.

We perform our observations on each user over a four-month time period, which consists of two months prior to treatment and two months post treatment. In order to form the treated group for each venue category, for the treatment of embarking on an offline activity (e.g., starting to go to bars), we identify and select those users who did not have any check-ins in the related venue category for a period of two months and subsequently made at least sixteen evenly distributed check-ins in the subsequent two months (i.e., at least one check-in every week, in order to be sure that they checked in at relevant venues consistently during the whole two month period). So, for instance, the treated group for ‘Embarking Bar’ would consist of those users who, based on their Foursquare check-ins, had not been to any venue in the Bar venue category for two months and then subsequently started to visit such venues at least once a week for two months.

A similar approach was adopted for abandoning treatments. In these treatments, we ensured that the subjects selected for the treated group had consistently checked in at venues of the given venue category at least sixteen times and at least once every week over a two-month period after which the subject would abandon venues in that venue category and would not make any check-ins in such venues for a two-month period. We report the number of users that were selected for our study in Table 3.

Table 3
The number of users in the study in both the control and treated groups. The number of users in treated and control groups are matched evenly and shown inside the parentheses.

	Embarking	Abandoning
Bar	916 (458)	650 (325)
Restaurant	662 (331)	514 (257)
Shopping	528 (264)	328 (164)
Sport	892 (446)	536 (268)

4.3. Control groups

The purpose of shaping the control group is to compare it with the treated group in order to show that the differences in the outcome are due to the exposure to the treatment and not due to other variables.

We used two approaches to form two different control groups, as follow:

Control group 1: We performed Propensity Score Matching (PSM) (Caliendo & Kopeinig, 2008) in order to rule out the impact of the confounding variables (Table 4) that, if not handled, could potentially yield unreliable causal relations. Following the PSM method, users in the treated and control groups were matched based on their propensity scores. Propensity scores are used for balancing the treated and control groups such that the distributions of the measured confounding variables are similar in the treated and the control sub-populations (Austin, 2011). The propensity score was obtained from a logistic regression fitted to predict the probability of a user being assigned to the treated group given the set of observed confounding variables of Table 4. We matched each user from the treated group with a user in the control group based on the nearest neighborhood of propensity scores. From among the nearest neighbors, the one that had the closest measurement on the dependent variable, prior to treatment, to the treated user's dependent variable, was selected for matching.

In Table 4, we report the standardized mean difference (SMD) of the confounding variables in order to show that the users in the treated group and the control group are *well balanced*. SMD is a measure of the difference of the averages of a confounding variable in the treated and control groups divided by the standard deviation of the treated group. It is generally suggested that absolute SMD values should stay below 0.25 (Rosenbaum & Rubin, 1985; Stuart, 2010) As shown in Table 4, SMD values for the confounding variables are within the acceptable margin. Given all the observed variables in Table 4 are in the acceptable range, the natural experiment can be considered to be random and hence, as mentioned by Althoff et al. (2017), other unobserved variables are possibly well balanced and not confounding the findings.

Control group 2: When using the PSM method, it would be ideal to control for all potentially confounding variables such as age and geographical location. However, such information is neither reliably nor comprehensively available on Twitter or Foursquare. In order to eliminate all demographic confounding variables, we match every user in the treated group in the treated year with herself but in a non-treated year. As an illustration of this, Fig. 1 shows the behavior of user @janedoe for a two-month period (April and May 2016), during which she does not go to a bar and subsequently starts to frequent at bars in June and July 2016. In order to shape the second control group, we match this user to herself but in a different year from when the treatment happened. In particular, we compare user @janedoe in April, May, June, July of 2016 with herself but in April, May, June and July of 2015. By doing so, we can eliminate all demographic confounding variables because the matched users in both groups are exactly the same. The impact of time-of-year has been eliminated, as well, because the outcome of both groups are for the same time of the year.

The employment of two control groups, one based on propensity score matching and the other based on matching the user with herself but in a non-treated time period, allows us to generalize our findings and eliminate any confounding variables. If the same degree of causality is observed based on both control groups, then the causal conclusion can be accepted with a higher confidence in the assumption that confounding variables were being properly controlled.

4.4. Outcome

The last element of the experimental design pertains to the measurement of the outcome. The outcomes that we are interested in are *social alignment* and *social convergence* of the users. As introduced in RQ1, we would like to study social alignment in terms of

Table 4
Balancing treated & control groups using PSM.

Variable	SMD
Number of Tweets	0.1947
Number of Twitter Followers	0.0327
Number of Twitter Friends	0.0916
Number of Check-ins	0.1903
Gender	0.12
Median Absolute SMD	0.12

aligning with public topical interests and sentiments. Additionally, as mentioned in RQ2, we explore the degree of changes of users' social alignment. To this end, we introduce four different metrics:

1 Social Topic Interest Alignment (*STIA*): how close are the user's topical interests to the topics of interest of the general community;
2 Social Topic Sentiment Alignment (*STSA*): how close the sentiments of a user and the overall sentiments of the general community are with regards to each active topic;
3 Social Topic Interest Convergence (*STIC*): how did the user's interest alignment change as a result of a treatment;
4 Social Topic Sentiment Convergence (*STSC*): how did the user's sentiment alignment change as a result of a treatment.

We formalize these four outcomes in the following section.

4.5. The measurement model

While the treatment is based on users' check-ins on Foursquare, the measurement of its outcomes is based on the content that the same users were posting on Twitter. We formalize a measurement metric for each of the four outcomes. In this section, we first define the preliminaries, and after that, we formalize the four metrics.

4.5.1. Preliminaries

The central theme of the four outcomes are the *topics* that are discussed on Twitter in a given time period; we refer to these as *active topics*. We are interested in understanding how the user's interests in these topics align with the community's interests and whether there is an alignment, or a lack thereof, in terms of sentiments expressed about these topics. We are further interested in knowing whether treatment can change the degree of alignment. For this reason, we formally define what our representation of a topic is and how a set of topics can be derived from the social content:

Definition 1. (Topic) Let M be the set of posts collectively published by all users and W, enumerated from 1 to N, be the set of all words. Topic z is a vector of N weights, i.e., $(q_z(w_1), \ldots, q_z(w_N))$ where $q_z(w_j)$shows the participation score of the j^{th} word in forming topic z.

In order to learn a set of topics (Z) from the content posted on Twitter, we employ TwitterLDA (Zhao et al., 2011), which is a topic modeling approach specifically developed for Twitter content. This method assumes a generative process where the user first selects a topic based on which a collection of words are drawn from the topic-word distribution to form a tweet. In order to train a topic model, we divide the users' tweets into continuous monthly time intervals each representing one document. We learn a topic model over all these documents across all time intervals. The learnt topic model will produce a set of topics, the distribution of words in each topic and the distribution of topics over tweets. It should be noted that the assumption in Zhao et al. (2011) is that each tweet is generated from a single topic and as such, each tweet is only assigned to one topic.

Based on the extracted topics from the content of the tweets, we let $Z = \{z_1, z_2, \ldots, z_K\}$ be K active topics on Twitter and define the User Topical Interest Profile ($UTIP_u^t$) and User Topical Sentiment Profile ($UTSP_u^t$) for each user u in time interval t as follows:

Definition 2. (User Topical Interest Profile) The interest profile of user $u \in U$ in time interval t, denoted as $UTIP_u^t$, is represented by a vector of weights over K topics, i.e., $(f_u^t(\mathrm{z}_1), \ldots, f_u^t(\mathrm{z}_K))$ where $f_u^t(\mathrm{z}_i)$ denotes the degree of u's interest in topic $z_i \in Z$ in time interval t. A user interest profile is normalized by $\ell 1$-norm. We define θ_m^z to be a binary variable that is 1 if tweet m belongs to topic z and 0 otherwise. On this basis, $f_u^t(z)$ is defined as:

$$f_u^t(z) = \frac{\sum_{m \in M_u^t} \theta_m^z}{|M_u^t|} \tag{1}$$

where M_u^t is the set of tweets posted by user u in time interval t .

We also formalize the User Topical Sentiment Profile based on the analysis offered by Linguistic Inquiry and Word Count (LIWC) (Paul & Dredze, 2011). LIWC is a widely-used text analysis program that counts the words in different categories such as various linguistic dimensions, grammatical structures, and social and psychological processes. In order to find sentiment distributions, we have used the positive and negative sentiments from the *psychological processes* category of LIWC.

Definition 3. (User Topical Sentiment Profile) The sentiment profile of user $u \in U$ in time interval t, denoted as $UTSP_u^t$, is represented by a vector of weights over K topics, i.e., $(g_u^t(\mathrm{z}_1), \ldots, g_u^t(\mathrm{z}_K))$ where $g_u^t(\mathrm{z}_i)$ denotes the average sentiment of user u with respect to topic $z_i \in Z$ in time interval t and is normalized by $\ell 1$-norm.

We measure the topical sentiment for each user u in time interval t for topic z based on the sentiment of the tweets published by u in t on topic z. As such $g_u^t(z)$ in Definition 3 is formalized as:

$$g_u^t(z) = \frac{\sum_{m \in M_u^t[z]} sentiment(m)}{\sum_{m \in M_u^t} sentiment(m)} \tag{2}$$

We have assumed, as suggested in Kramer (2010), that the differences between the positive sentiment and negative sentiment associated with a tweet m by LIWC produces a value expressed by *sentiment(m)* for tweet m. $M_u^t[z]$ represents a subset of M_u^t that is related to topic z.

To calculate social alignment of users to public interests and sentiments, we would need to also determine how the general

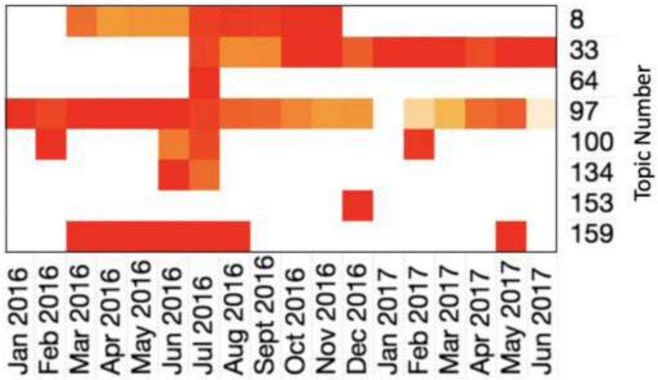

Fig. 3. A visual representation of the community topical profile for a selected set of topics shown in Table 5 (Red indicates high interest while White shows no interest from the community). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

community views the active topics in terms of interest and sentiment. To this end, we extend Definitions 2 and 3 as follows:

Definition 4. (Community Topical Interest Profile) Let U denote a set of users. The community topical interest profile, denoted by $CTIP^t$, is represented by a vector of weights over K topics, i.e., $(h^t(z_1), \ldots, h^t(z_K))$ and is normalized by $\ell1$-norm. $h^t(z)$ is defined as:

$$h^t(z) = \frac{\sum_{m \in M^t} \theta_m^z}{|M^t|} \tag{3}$$

where M^t is the set of tweets posted by all users U in time interval t. As such $CTIP^t$ represents the normalized topic distribution for all tweets published in time interval t. Based on the information captured in$h^t(z)$ over several time intervals (t), it is possible to contrast it against the interest profile of each user u to measure how much the user is aligned with the interests of the general public.

For the sake of better understanding, Fig. 3 visualizes a representation of the community topical interest profile for each month of our dataset using a heatmap for a selected set of topics shown in Table 5. The heatmap visualizes $h^t(z)$as defined in Definition 4 for different time intervals (t) and shows how the topical interests of the general community have changed over the examined time intervals. Table 5 lists the top-10 words for each topic illustrated in Fig. 3. As shown in the figure, Topic 8 emerged in March 2016, and started to fade in popularity in November 2016 when people started to talk more about the US elections (Topic 33). Topic 100 is related to the Grammy Awards and was popular only in a few months of the examined time period. Topic 97, related to New York city, was popular all year round. Topic 134 is related to Brexit. Brexit happened in June 2016 and people in the UK voted to leave the EU. Thus, it is not surprising that this topic was popular in June and July 2016 (Fig. 3).

We also define a *community sentiment profile* to show the sentiments of the general community towards the set of active topics Z.

Definition 5. (Community Topical Sentiment Profile) The community sentiment profile in time interval t, denoted as $CTSP^t$, is represented by a vector of weights over K topics, i.e., $(v_t(z_1), \ldots, v_t(z_K))$ where $v_t(z_i)$ is computed as the average sentiment of users with respect to topic $z_i \in Z$ and is normalized by $\ell1$-norm. $v_t(z)$is defined as:

$$v^t(z) = \frac{\sum_{m \in M^t[z]} sentiment(m)}{\sum_{m \in M^t} sentiment(m)} \tag{4}$$

where $M^t[z]$ is a subset of M^t denoting tweets posted by all users in time interval$t$ about topic z.

Similar to Fig. 3, Fig. 4 is a heatmap but it visualizes the *Community Sentiment Profile* over time. The purpose of this figure is to show how people's sentiments about different topics change over time. For example, users' sentiment about Topic 8 in March 2016 was positive and it became even more positive in June and July that year. However, the sentiments associated with Topic 8 became negative over time and turned completely negative by November 2016. As another example, Topic 33 attracted completely negative sentiments over time. There are a few topics that attracted a lot of positive sentiments from the users, e.g., a topic related to Christmas (Topic 153) is one such example.

Table 5
Top-10 words of topics illustrated in Figs. 3 and 4.

Topic 8	Trump, vote, hillary, debate, clinton, hillaryclinton, election, donald, realdonaldtrump
Topic 33	Trump, realdonaldtrump, president, clinton, news, russia, media, fbi, election, america
Topic 64	Love, song, smule, singkaraoke, check, sheeran, cover, video, justinbieber, music
Topic 97	Ny, york, brooklyn, park, nyc, queens, square, city, subway, station
Topic 100	Grammys, song, ladygaga, show, album, music, video, beyonce, taylor, adele
Topic 134	Brexit, uk, people, eu, vote, bbc, news, london, election, britain
Topic 153	Christmas, year, happy, merry, thanksgiving, holiday, santa, eve, tree, xmas
Topic 159	Chicago, land, louis, saint, denver, park, starbucks, cta, restaurant, club

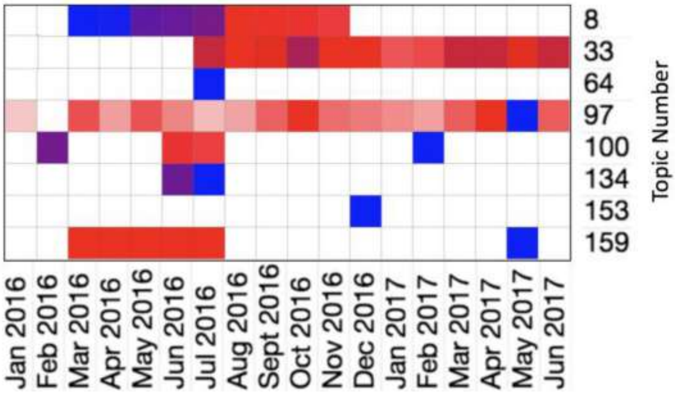

Fig. 4. The visual representation of the community sentiment profile for a selected set of social topics shown in Table 5 (Red and Blue indicate negative and positive sentiments, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

It is possible to contrast the trends observed in the community sentiment profile with the sentiment profile of each user to see if the user's sentiments are aligned with that of the general community. For instance, does the user show similar positive sentiment towards a topic that is generally liked by the community and vice versa.

4.5.2. Quantifiable metrics

We build on the four preliminary definitions, two at the user level and two at the community level, to formally quantify the four outcomes in our experiments. We first focus on *social alignment* outcomes:

1) *Social Topic Interest Alignment*, which is the degree of difference between a User Topical Interest Profile and the Community Topical Interest Profile, and
2) *Social Topic Sentiment Alignment*, which is the difference between a User Topical Sentiment Profile and the Community Topical Sentiment Profile.

Definition 6. (Social Topic Interest Alignment) The Social Topic Interest Alignment of user $u \in U$ in time interval t, denoted as $STIA_u^t$, is defined as the similarity between User Topical Interest Profile of user u in time interval t, $UTIP_u^t$, and Community Topical Interest Profile in the same time interval t, $CTIP^t$ as follows:

$$STIA_u^t = 1 - \left(\sum_{z \in Z} h^t(z) - f_u^t(z) \right) \tag{5}$$

This outcome shows how close the user and the community are in terms of the topics they are interested in. For example, in Fig. 1, in April and May, user @janedoe talks only about one popular topic (Topic 8). This suggests, a significant difference between her topical profile and the community topical profile. But, in June, she engages in 4 topics of general interest and in July she discusses 7 topics of interest to the overall community. This indicates that the user's interest profile is gradually becoming closer to the community's interest profile. If the topic distribution for user u is completely the same as the community's topic distribution, i.e., there is a perfect alignment between the user's and community's topical interests, the User Topical Profile and Community Topical Profile will be the same and hence STIA will be equal to one. On the other hand, if user u did not talk about any popular topics, *STIA* for the user will be zero, pointing to the fact that the user's topical interests are completely unaligned with the community.

Definition 7. (Social Topic Sentiment Alignment) The Social Topic Sentiment Alignment of user $u \in U$ in time interval t, denoted as $STSA_u^t$, is defined as the difference between Community Topic Sentiment Profile in time interval t, $CTSP^t$, and User Topic Sentiment Profile of user u in the same time interval, $UTSP_u^t$, as follows:

$$STSA_u^t = 1 - \left(\sum_{z \in Z} v^t(z) - g_u^t(z) \right) \tag{6}$$

This metric shows how close are a user's sentiments to the community's sentiments on active topics. For example, Fig. 5 shows the distribution of topic-related sentiments of user @janedoefor the months four, five, six, and seven of 2016. Red circles in the figure are showing @janedoe's sentiments for the top-10 topics, whereas blue dots represent sentiments expressed by the general public on the same topics. As can be observed in the figure, in April and May of 2016, @janedoe's sentiment was negative for Topic 8, while the public's sentiment about this topic was positive. In June and July of 2016, during which user @janedoe started to go to bars, her

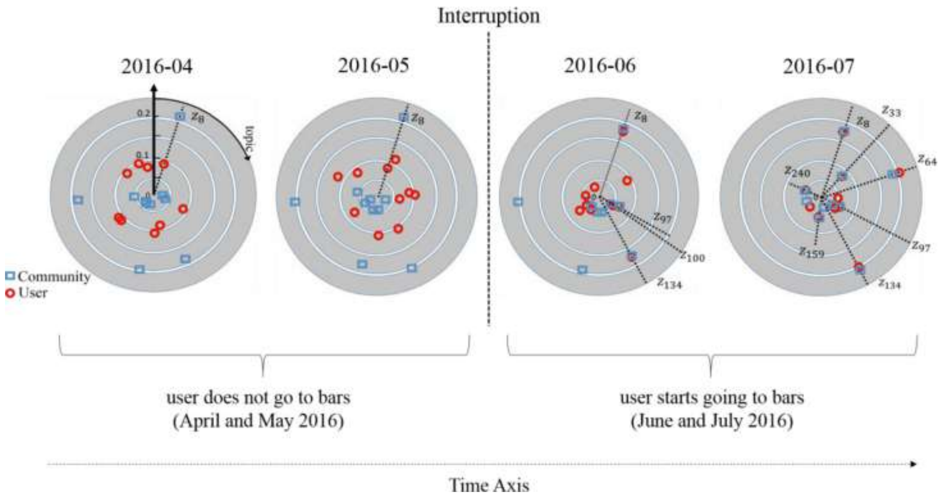

Fig. 5. Topical sentiment distribution of user @janedoe in contrast to the general public. Topics are placed around the radar and denoted by $\boldsymbol{z_i}$ and sentiment related to each topic is specified by distance from the center. For the sake of visualization, sentiment measurements are normalized to the range of [0,1] where zero indicates the most negative sentiment while one denotes the most positive.

sentiments changed and became closer to that of the general public. Based on Fig. 1 and Fig. 5, one can conclude that not only user @ janedoe started talking about similar topics as the community after she began frequenting bars, but also her sentiments became closer to those of the general public.

Now, we are further interested in understanding how users' behavioral patterns change over time. More specifically, we would like to investigate whether a user's topic and sentiment alignment change over time. As such, we track the changes on these metrics over consecutive time intervals and define two metrics as follows:

Definition 8. (Social Topic Interest Convergence): The Social Topic Interest Convergence of user $u \in U$ in time interval t, denoted as $STIC_u^t$, is a measure of interest change over two subsequent time intervals and is defined as the difference between Social Topic Interest Alignment of user u in time interval t and time interval $t-1$, shown as:

$$STIC_u^t = (STIA_u^t - STIA_u^{t-1}) \tag{7}$$

This metric allows for examining the change in a user's Social Topic Interest Alignment over time.

Definition 9. (Social Topic Sentiment Convergence): The Social Topic Sentiment Convergence of user $u \in U$ in time interval t, denoted as $STSC_u^t$, is a measure of change in a user's sentiments over two subsequent time intervals and is defined as the difference between Social Topic Sentiment Alignment of user u in time interval t and time interval $t-1$. Hence, the formula for Social Topic Sentiment Convergence is defined as:

$$STSC_u^t = (STSA_u^t - STSA_u^{t-1}) \tag{8}$$

This metric allows for examining changes in a user's Social Topic Sentiment Alignment over time as a result of some treatment or just based on progression of time.

In summary, Table 6 shows how the defined metrics correspond to our research questions and will be used to measure the expected outcomes in our study.

5. Results

In this section, we present our findings regarding the impact of different treatments (offline activities) on the dependent variables (metrics of the research questions, Table 6).

Table 6
The correspondence between research questions and the defined metrics (mapping to Table 1).

	Topical interests	Topical sentiments
Social alignment	RQ1.1 (Definition 6 - STIA)	RQ1.2 (Definition 7 - STSA)
Social convergence	RQ2.1 (Definition 8 - STIC)	RQ2.2 (Definition 9 - STSC)

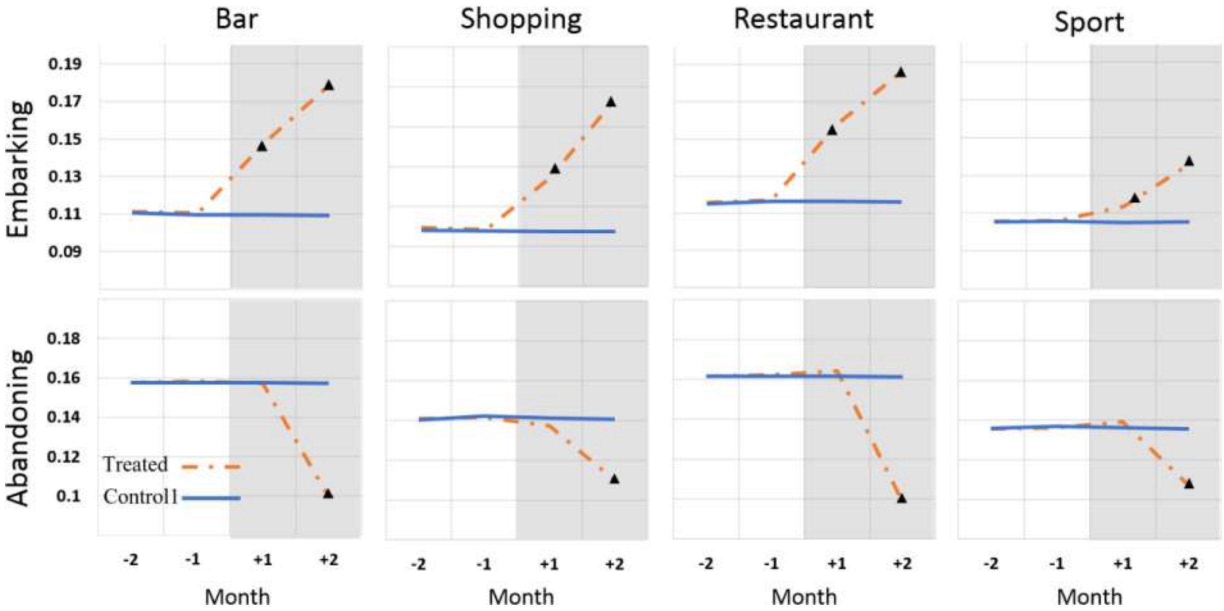

Fig. 6. Impact of socially situated offline activities on users' social topic interest alignment (*y*-axis). Statistical significance is noted by ▲ on the treated group.[4] The control group is based on PSM (Control Group 1). Months −1 and −2 are the months prior to the treatment (white background) and Months +1 and +2 are the months after the treatment (grey background).

5.1. RQ1.1: social topic interest alignment

We separated the involvement of users with offline activities into *embarking* and *abandoning* treatments and present our findings separately for each of the four venue categories introduced in Table 2, i.e., Bar, Shopping, Restaurant, and Sport. The observations for Social Topic Interest Alignment based on the treated group and the control group matched through PSM (Control Group 1) are reported in Fig. 6. As the figure indicates, when users embark on an offline activity that increases their exposure to or their retrieval from social situations, the alignment of their topical interests with the interests of the community increases. The reverse happens when users abandon such an offline activity, namely their alignment to the community with respect to the topical interests decreases. In all four venue categories and for the embarking treatment, the increase in topic alignment of the treated group is statistically significant compared to Control Group 1 in both months after the treatment (months +1 and +2). However, for the abandoning treatment, statistical significance is only observed in all four categories in month +2. This means that embarking on a socially situated activity results in an immediate and lasting impact on users' topic alignment, while abandoning such an activity gradually shifts users interests away from the community's interests.

To better examine the effect of offline activities, we compared users in the treated group with themselves but in a different time period, denoted as Control Group 2 (Section 4.3). The reason we compare these two groups is to understand whether the differences observed between the first control group and the treated group were due to confounding variables we may not have been able to control for in Control Group 1 (i.e., demographic factors or differences in the examined time period), or primarily due to the independent variables, i.e., the offline activities. We hypothesize that similar trends observed between the treated group and each of the two control groups indicate that the detected statistical differences are due to the independent variables. In Fig. 7, we can see the results of topic interest alignment for Control Group 2 compared to the Treated Group. This figure shows the same behavior trend as the one observed in Fig. 6. Therefore, considering that with the two control groups we have controlled for all potentially relevant confounding factors, we can conclude that the observed statistical differences are primarily due to the *treatment*.

Finding 1. Those offline activities that impact the exposure of users to or their retrieval from social situations impact how users topical interests align with the community. Embarking on an offline activity that exposes a user to social situations has *immediate and steady* positive impact in terms of aligning user's interests with the topics of interest to the general community whereas abandoning such an offline activity has a delayed negative impact on topic interest alignment.

[4] All statistical significance results are based on paired *t*-test at 0.05.

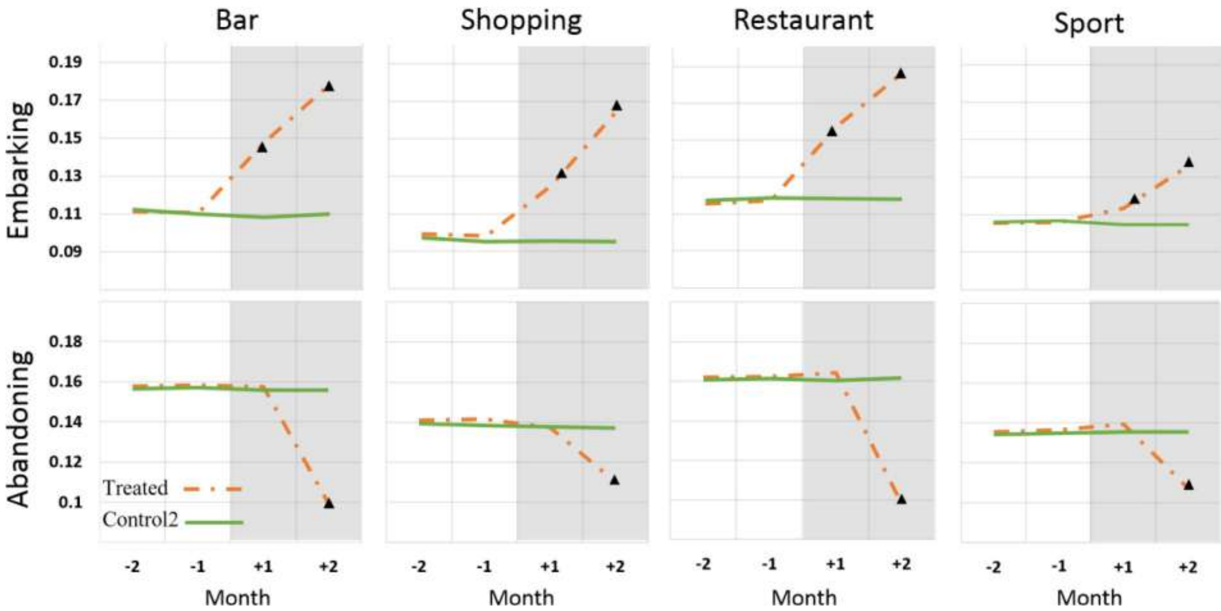

Fig. 7. Social topic interest alignment for Treated Group and Control Group 2. Labels on the two axes have the same meaning as on Fig. 6.

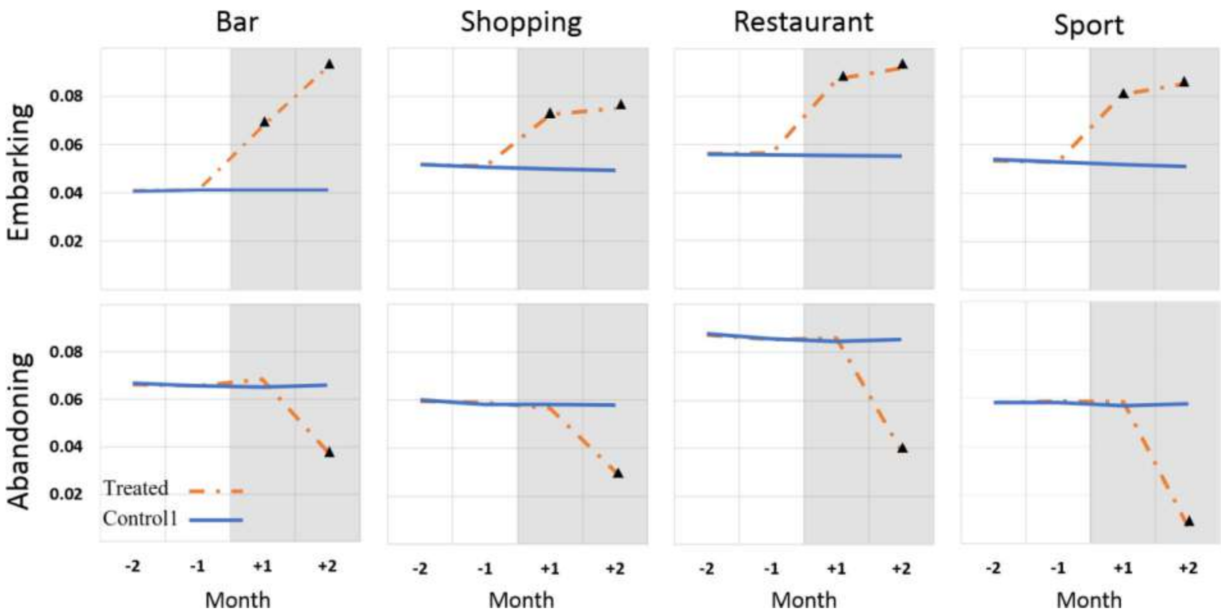

Fig. 8. The impact of socially situated offline activities on user's Social Topic Sentiment Alignment (y-axis). ▲ denotes statistical significance. The control group is based on PSM (Control Group 1). Months −1 and −2 are the months prior to treatment (shown in white) and Months +1 and +2 are the months after the treatment (shown in grey).

5.2. RQ1.2: social topic sentiment alignment

Similar to Social Topic Interest Alignment, embarking and abandoning treatments have inverse impact on Social Topic Sentiment Alignment, as shown on Fig. 8. Specifically, the figure suggests that embarking on a socially situated offline activity increases social sentiment alignment with respect to the active topics. In other words, when a user embarks on such an offline activity, it becomes more likely for that user to share sentiments of the general community regarding the active topics. On the other hand, when a user abandons a socially situated offline activity, it becomes less likely that the user will adopt a sentiment towards an active topic that will be inline with the overall sentiment expressed by the general community.

In Fig. 9, social topic sentiment alignment for the treated group and the second control group is shown. The figure shows a trend similar to the one on Fig. 8. This indicates that the observed impact is primarily due to the treatments (embarking or abandoning a

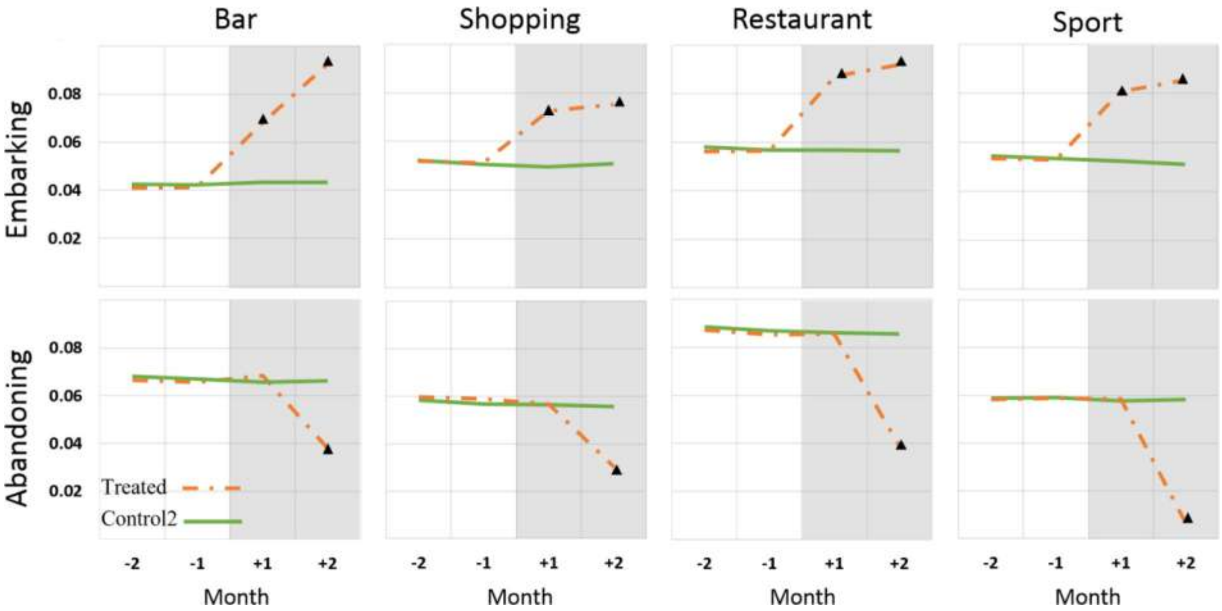

Fig. 9. Social topic sentiment alignment for Treated Group and Control Group 2. Months −1 and −2 are the months prior to treatment (shown in white) and Months +1 and +2 are the months after the treatment (shown in grey).

socially situated offline activity) and other confounding variables that may have not been controlled for in Control Group 1 do not play a role in the observed statistical difference.

> **Finding 2.** An offline activity that alters users' exposure to or their retrieval from social situations can provide *significant* measurable influence on the alignment of users' topic sentiments with that of the community. The influence is often observed immediately after a user embarks on a new offline activity that places him/her in a social context and increases the likelihood of the user sharing the sentiment of the general community regarding active topics. On the other hand, the impact is observed in a delayed fashion when the user abandons such an offline activity and takes the form of the user's sentiment *diverging* from the sentiment of the general community on active topics.

5.3. RQ 2.1: social topic interest convergence

Our observations with regards to RQ 2.1, based on the comparison of the treated group and the control group formed using PSM (Control Group 1) are depicted in Fig. 10. The figure shows that both embarking and abandoning activities for all venue categories have statistically significant impact on the degree of users' social topic interest convergence. Given social topic interest convergence is, in essence, the difference in a user's social topic interest alignment between two consecutive time periods, the findings indicate that users are highly likely to change their social topic interests as a result of an offline activity that places them in a social situation. For users embarking on such an offline activity, the impact is present in the first month after the treatment (Month +1) for all categories of venues except the sport-related ones; only for sport activities the effect is delayed and occurs in the second month after the treatment (Month +2). In contrast, users who abandon a socially situated offline activity have a lengthier window of impact where the degree of convergence is impacted in Month +2 (two months after the treatment).

Fig. 11 contrasts the values of social topic interest convergence for the treated group to that of the second control group. The figure shows a very similar trend for the control group as the one shown on Fig. 10. In particular, the social topic interest convergence for all months of the second control group are either zero or very close to zero. This suggests that when users are not exposed to any socially situated offline activity their topical interests do not change over time. On the other hand, the social topic interest convergence of the treated group changes after receiving the treatment. Given the two groups consist of the same set of users but in different time intervals, with the only difference that the users were exposed to the treatment in one time interval and not in the other, this reinforces our findings in Fig. 10, indicating that the change observed on social topic interest convergence of the treated group can be attributed to the treatment and not other possibly uncontrolled confounding variables.

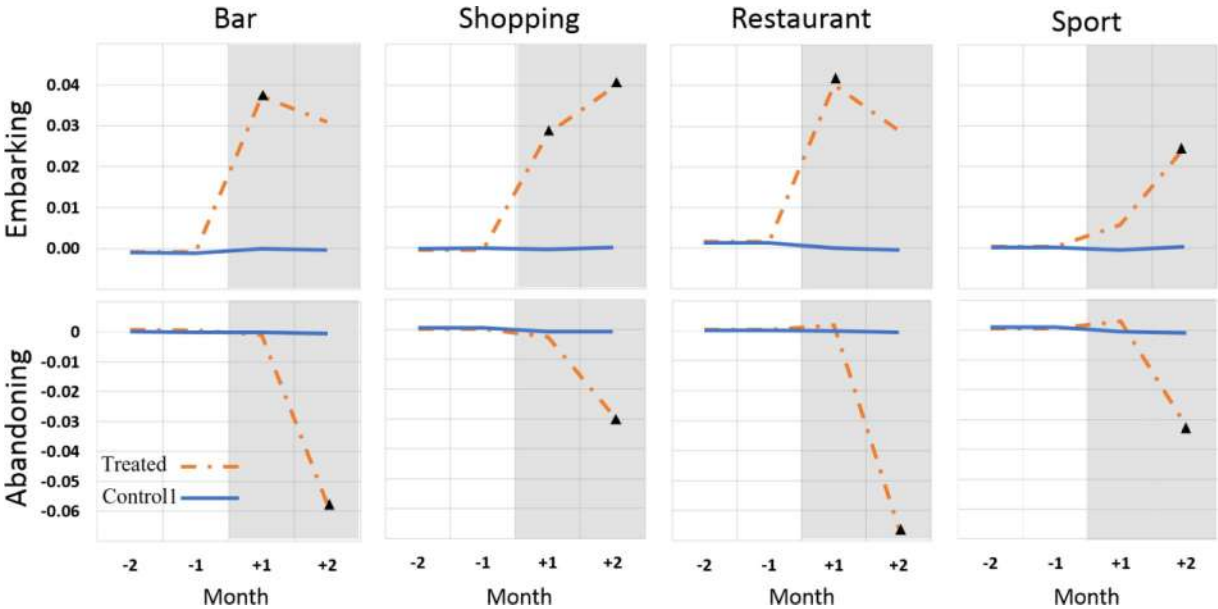

Fig. 10. Impact of socially situated offline activities on user's social topic interest convergence (y-axis). ▲ shows statistical significance. The control group is based on PSM (Control Group 1). Months −1 and −2 are the months prior to treatment (shown in white) and Months +1 and +2 are the months after the treatment (shown in grey).

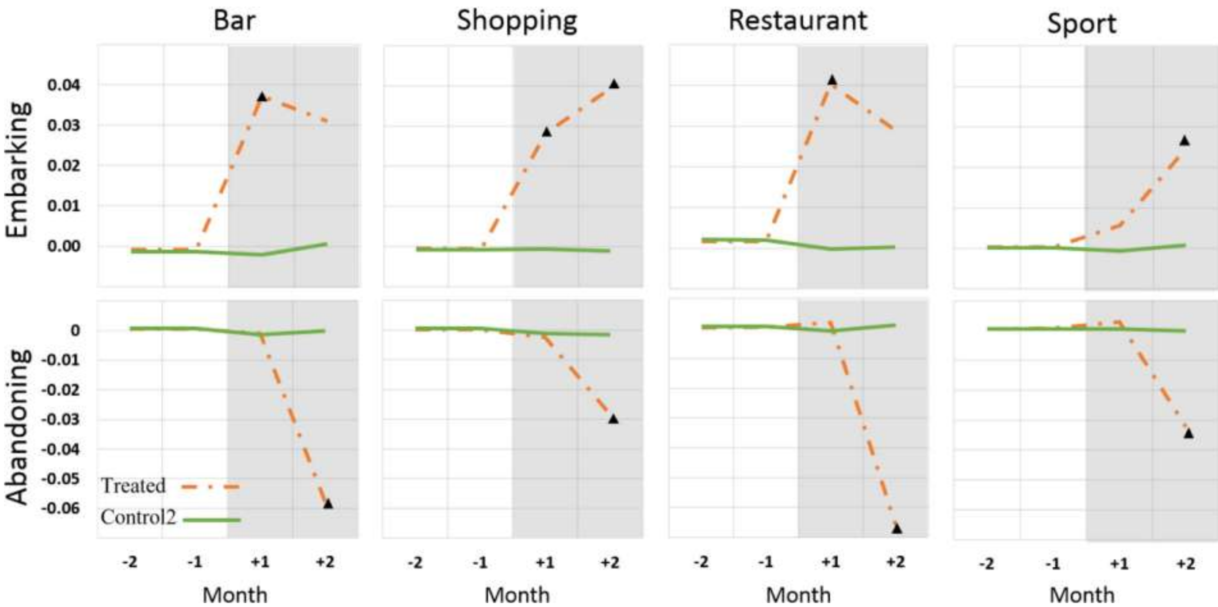

Fig. 11. Social topic interest convergence for Treated Group and Control Group 2. Months −1 and −2 are the months prior to treatment (shown in white) and Months +1 and +2 are the months after the treatment (shown in grey).

Finding 3. Initiating or discontinuing socially situated offline activities, which constitute *embarking* and *abandoning* treatments, lead to changes in social topic interest *convergence* and point to situations when users are most likely to change their topic interests. Users exposed to the *abandoning* treatment have a lengthier *window of impact* as the impact on their interests is not immediately observed in Month +1 compared to the *embarking* treatments that show an immediate impact in Month +1 (except for the Sport category).

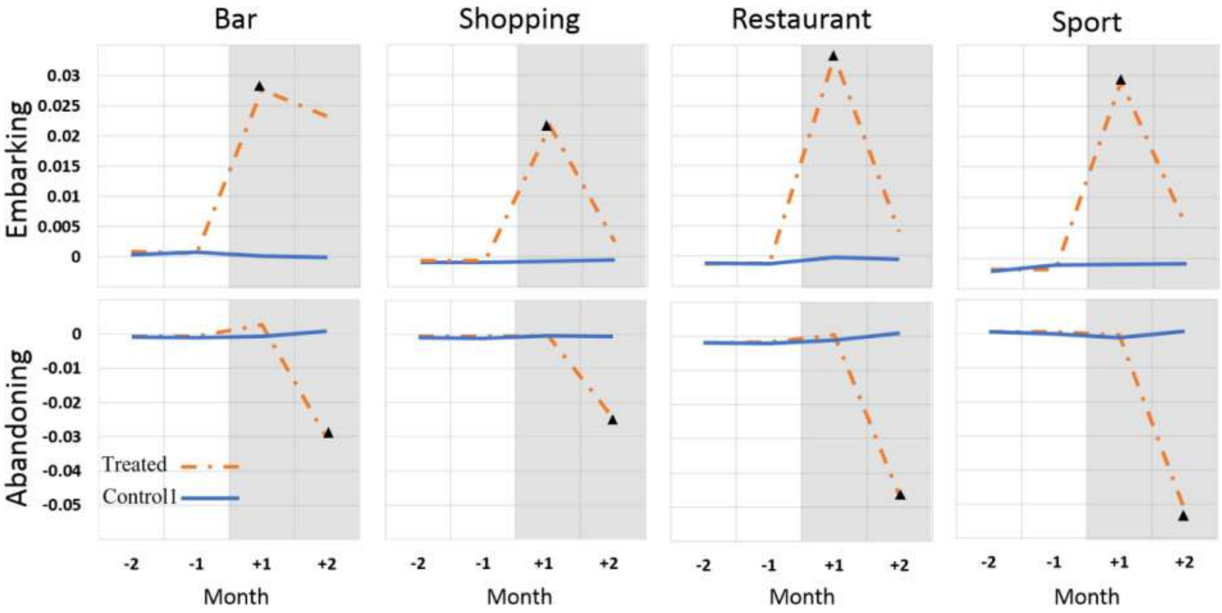

Fig. 12. Impact of socially situated offline activities on user's social topic sentiment convergence (*y*-axis). ▲ reports statistical significance. The control group is based on PSM (Control Group 1). Months −1 and −2 are the months prior to treatment (shown in white) and Months +1 and +2 are the months after the treatment (shown in grey).

5.4. RQ 2.2: social topic sentiment convergence

Fig. 12 shows how the difference in topic sentiments between users and the community changed during two months prior to the treatment and two months after, based on the data for the treated group and the PSM-based control group (Control Group 1). As shown on the figure, when a person embarks on an offline activity that increases his/her exposure to or retrieval from social situations, a significant change occurs in the first month after the treatment. However, the difference decreases in the second month, suggesting that embarking on such an offline activity has an immediate effect on users' sentiments towards active social topics and then maintains this impact in a subsequent month (by not changing from Month +1). In other words, the users who starta socially situated offline activity tend to align their topic sentiments with that of the general community in the first month and then maintain

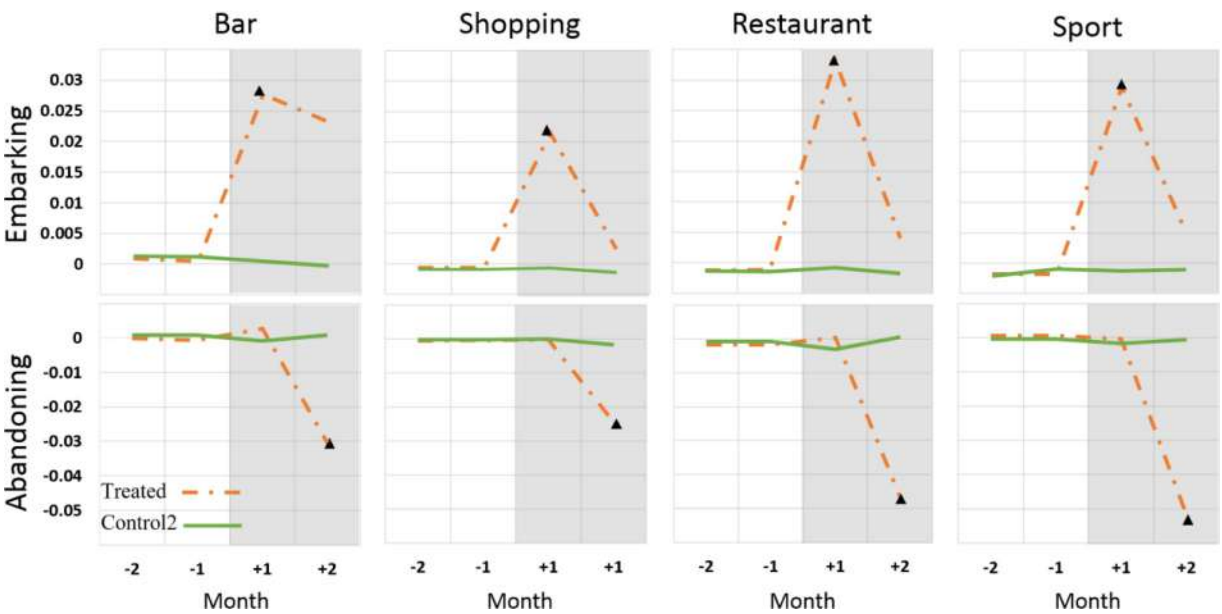

Fig. 13. Social topic sentiment convergence for Treated Group and Control Group 2. Months −1 and −2 are the months prior to treatment (shown in white) and Months +1 and +2 are the months after the treatment (shown in grey).

these newly adopted sentiments going forward in the second month. On the other hand, those users who leave a socially situated offline activity tend to gradually shift away from the sentiments shared by the community and the difference becomes statistically significant in the second month after the treatment.

Fig. 13 shows the social topic sentiment convergence for the treated group compared to the second control group. From this figure, we can see that while both the treated group and the second control group consist of the same set of users (in two different years), the changes observed after Month +1 for the embarking treatment and after Month +2 for the abandoning treatment are only observed in the treated group, which further means that users do not normally experience much change in their sentiments about active topics if they are not exposed to some external treatment.

Finding 4. Abandoning a socially situated offline activity can cause *delayed yet significant* impact on users' sentiments towards active topics, leading them away from the sentiments expressed by the general public. On the other hand, embarking on a socially situated offline activity will cause the users to move their sentiments on active topics close to the general public and remain as such when moving forward.

6. Predicting behavior change

Having observed causal impacts between users' socially situated offline activities and their online behavior, we considered it reasonable to expect that a predictive model based on the examined kinds of offline activities as independent variables would outperform models built using variables identified as potentially relevant through correlation or otherwise. In other words, independent variables that have causal impact on the outcome should have a stronger predictive power for determining the outcome than variables identified e.g. through correlational analysis. To examine the validity of this assumption, we adopted the strategy suggested in Althoff et al. (2017) to design the prediction task as a binary classification problem where the goal was to predict whether the four dependent variables would increase or decrease solely based on the information about the kinds of offline activities examined in our study. More specifically, we have built two competing models: (i) one based on the variables introduced in Table 4 and (ii) the other based on eight features each denoting whether the user either embarked on or abandoned one of the four offline activities (associated with the four kinds of venues listed in Table 2).

We used a random split of 80/20 for training and testing and built a Gradient Boosted Tree (GBT) (Chen & Guestrin, 2016) model with the default parameter settings of scikit-learn.[3] As the baseline, we trained a similar GBT model based on the features introduced in Table 4. In Table 7, we report the percentage of improvement achieved over the baseline based on classification accuracy as the performance measure. As shown in Table 7, the examined offline activities, when used as features, can be strong predictors of online behavior. In line with the findings of our causal study, we find here that the performance of the predictive model is weaker on sentiment-based dependent variables and stronger on interest-based features. Furthermore, prediction accuracy improvement is consistent across the two months following the intervention (months +1 and +2).

Overall, we show that it is possible to predict how users will behave online by knowing their offline activities and the prediction accuracy is significantly higher when based solely on the users' offline behavior compared to all the variables that were controlled for as described in Table 4. This means that we can predict how a user will behave in an online social platform, in terms of reacting (both with regards to interest and sentiment) to active online topics. This can be used, for example, in marketing for the purposes such as *information customization, recommendation,* and *persuasion*. For instance, let us assume that a user such as 'John Doe' has just started going to the *Fit Factory Fitness* sports club in Toronto. With this information, a marketer would know that this user is very likely to be impacted by normative social influence and hence his topical and sentimental interests are likely to gradually converge towards the topical and sentimental interests of the community. As such, the products with the highest likelihood of being of interest to 'John Doe' would be those that are widely discussed by and are of interest to the users who already go to the *Fit Factory Fitness* club. Such products could be identified by reviewing the publicly shared social content, e.g. tweets, of those users who have also checked in at the *Fit Factory Fitness* club venue on Foursquare. Therefore, the marketer can directly reach out to 'John Doe' with offers to sell or provide these products to him.

7. Findings and threats to validity

The central objective of this paper has been to explore whether the reflection of users' socially situated offline activities can be observed in their online behavior. We specifically explored whether embarking on or abandoning offline activities that can expose a user to or retrieve the user from social situations lead to changes in the person's online behavior. Based on a *natural experiment,* we found that:

1 A change in the user's offline behavior that i impacts the level of his/her exposure to social situations can cause a change in the user's online topical interests and sentiment. This is an interesting observation in that it shows that socially situated offline activities can have statistically significant impact on users' online behavior.

[3] http://scikit-learn.org/.

Table 7
Percentage of improvement (Δ%) when predicting the dependent variables based on the examined offline activities as the only features compared to the features introduced in Table 4.

	Dependent variable	Month +1	Month +2
1	Social Topic Interest Alignment	43.90%	62.05%
2	Social Topic Sentiment Alignment	22.23%	18.52%
3	Social Topic Interest Convergence	60.80%	57.01%
4	Social Topic Sentiment Convergence	29.56%	32.86%

2 Our study has demonstrated that users' topical interests and sentiments converge in a statistically significant manner towards interests and sentiments of the general public when the users embarked on or abandoned socially situated offline activities while all other potentially relevant variables were controlled for. This finding validates our hypothesis that a change in users' offline behavior that affects the level of their exposure to social situations can impact the users topical and sentimental expressions, as expected based on the theory of normative social influence.
3 The causal observations between the examined offline activities and the users' online social behavior can be used to build effective predictive models of users' online behavior. We have shown that by only considering users' offline activities it is possible to accurately predict users' online topical interests and sentiments.

The findings of our paper are complementary to the earlier findings by other researchers such as Althoff et al. (2017) who found that online behavior reflects itself in people's offline behavior. Althoff et al. argued that such a causal impact can be used within the context of persuasive technology to bring about positive behavioral change to the users, e.g., increasing their physical and exercise activities. Our work shows that the reverse of this process is also possible where the knowledge of the users' offline activities can be applied to bring about positive behavioral change in an online social network.

It is important to acknowledge the threats to the validity and generalizability of this work from the perspectives of both internal and external validity. Internal validity is primarily concerned with the extent to which the presented evidence support the identified causal relation. A major threat to internal validity is *confounding effect,* which refers to cases where the changes in the dependent variable can be attributed to a third unobserved confounding variable. In our work, we have carefully employed two control groups, one based on propensity score matching and one based on matching the users with themselves. While the first control group ensures that the distribution of users in the treatment and control groups are balanced in terms of social network attributes such as number of followers, friends, and check-ins, the second control group matches each user with itself but at a different time, as such potentially eliminating the impact of user-specific psychological and social characteristics. It is however possible, yet not likely, that some external unobserved confounding variable such as a real world event may have confounded the findings. The reason we believe this may be less likely is because the observations made across control groups 1 and 2 are quite consistent while the time difference of one year exists between the two control groups; hence reducing the possibility for such a confounding variable.

Furthermore, external validity is concerned with the generalizability of the findings beyond the scope of the experimental setting. We believe that in our work threats to external validity need to be considered from a *population validity* perspective. It must be mentioned that given our work requires the users to have been active both on Twitter and Foursquare, the findings can only be generalized to the population of users who have joined both platforms. In other words, user characteristics, such as social awkwardness and persuasibility, may be different between users who are only active on Twitter or Foursquare and those who are active on both. Given the fact that studies have shown that users with similar characteristics are active on both social networks (Han, Wang, Crespi, Park & Cuevas, 2015), it can be assumed, with reservation, that the findings can be generalized to the population of users who are active across a range of online social networks. Furthermore, we have investigated the behavior of users based on four Foursquare venue categories, which can be considered a limiting aspect with regards to population validity. We additionally explored the possibility of including other Foursquare category venues such as *Travel*, e.g., airports and tourist information centers, *Beauty*, e.g., tanning salon and tattoo parlor, and *Education*, e.g., schools and student centers. However, we were not successful in identifying a sufficient number of users and check-ins in these venue categories to warrant reliable analysis of the data. Moreover, while this paper has explored offline activities that place or retrieve users from social situations, it would also be highly interesting to see how offline activities that do not have social exposure impact the users' online activities. In this current study, we were not able to perform such study given the nature of Foursquare as a social app, which primarily promotes check-ins at social venues. As such, we did not have access to sufficient data on other types of venues to be able to undertake this objective.

The final aspect of external validity could be related to our inclusion criteria for the users in the treated group. The inclusion criteria specified as eligible those users who have made 16 checkins over a period of two months with at least one check-in every week. There were methodological reasons for this condition. In particular, given that we were interested in the involvement of users with offline activities, we wanted to ensure that only those users are selected who were really involved with offline activities. As such, we defined involvement with an activity as having at least one check-in in each week. We also considered selecting users with at least one check-in per day. However, this criterion did not produce sufficient number of users in the treated group and as such was not further explored. The other option was to constrain the selection to only those users who had made at least one check-in per month and at least 16 check-ins overall. The issue we faced was highly disproportionate distribution of check-ins when such condition was used. More specifically, a significant number of users who were included in the treated group based on this condition had one check-in in one month and a high number of check-ins, focused around few days, in the other month. We did not consider such sporadic behavior to be representative of involvement with an offline activity and hence did not pursue it in our selection of the treated group. It should be noted that our findings are generalizable to the extent of such selection condition.

8. Concluding remarks

We have reported on an extensive natural experiment that examined the impact of users' offline activities, which impact the level of users exposure to social situations, on their online behavior, in particular topical interest and topical sentiment aspects of the users' online behavior. A unique aspect of our work is that it benefits from a curated dataset that aligns users activities from Twitter, for capturing online user behavior, and Foursquare, to obtain offline user activity information. Our findings indicate that socially situated offline activities can direct users towards or away from the general public's topical interests and sentiment, depending on whether the user is embarking on or abandoning such offline activity. Given the findings of this paper, our work in the future will explore the following directions:

1 There has already been extensive work in sociology showing that a social phenomenon can have contagious impact. For instance, Christakis and Fowler (Christakis & Fowler, 2013) have shown this contagious effect for a number of characteristics such as obesity and smoking. We are interested in exploring whether users' offline activities have any impact on topical and sentiment profiles of the users' social connections. In other words, can the indirect impact of peoples' offline activities be observed in the online behavior of their social connections.
2 Furthermore, while this paper has explored topical and sentiment profiles of users, language analysis software, such as LIWC, do provide deeper insight into other social and psychological processes such as anxiety and depression. We are interested in performing more in-depth analysis of these more detailed psychological processes to see whether similar causal relations can be observed.

Acknowledgements

The authors graciously acknowledge funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) - RGPIN-2015-06118.

References

Althoff, T., Clark, K., & Leskovec, J. (2016). Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *Transactions of the Association for Computational Linguistics, 4*, 463–476. https://transacl.org/ojs/index.php/tacl/article/view/802.
Althoff, T., Jindal, P., & Leskovec, J. (2017). Online actions with offline Impact: How online social networks influence online and offline user behavior. *Proceedings of the tenth ACM international conference on web search and data mining (WSDM '17)* (pp. 537–546). New York, NY: ACM. https://doi.org/10.1145/3018661.3018672.
Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*(3), 399–424.
Bagozzi, R. P., Gopinath, M., & Nyer, P. U. (1999). The role of emotions in marketing. *Journal of the Academy of Marketing Science, 27*(2), 184–206.
Bandura, A. (1977). *Social learning theory.* New York: General Learning Press.
Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31–72.
Celli, F., Ghosh, A., Alam, F., & Riccardi, G. (2016). In the mood for sharing contents: Emotions, personality and interaction styles in the diffusion of news. *Information Processing & Management, 52*(1), 93–98.
Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (KDD '16)* (pp. 785–794). New York, NY: ACM. https://doi.org/10.1145/2939672.2939785.
Choudhury, M. D., Counts, S., & Horvitz, E. (2013a). Predicting postpartum changes in emotion and behavior via social media. *2013 ACM SIGCHI conference on human factors in computing systems, CHI '13, Paris, France, April 27–May 2, 2013* (pp. 3267–3276). . https://doi.org/10.1145/2470654.2466447.
Choudhury, M. D., Gamon, M., Counts, S., & Horvitz, E. (2013b). Predicting depression via social media. *Proceedings of the seventh international conference on weblogs and social media, ICWSM 2013* (pp. 1–10). . http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6124.
Choudhury, M. D., & Kiciman, E. (2017). The language of social support in social media and its effect on suicidal ideation risk. *Proceedings of the eleventh international conference on web and social media, ICWSM 2017* (pp. 32–41). . https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15662.
Choudhury, M. D., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2016a). Discovering shifts to suicidal ideation from mental health content in social media. *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2098–2110). . https://doi.org/10.1145/2858036.2858207.
Choudhury, M. D., Sharma, S. S., & Kiciman, E. (2016b). Characterizing dietary choices, nutrition, and language in food deserts via social media. *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing, CSCW 2016* (pp. 1155–1168). . https://doi.org/10.1145/2818048.2819956.
Christakis, N. A., & Fowler, J. H. (2013). Social contagion theory: examining dynamic social networks and human behavior. *Statistics in Medicine, 32*(4), 556–577 2013.
Cialdini Robert, B. (1993). *Influence: Science and practice* (3rd ed). New York, NY: Harper.
Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology, 58*(6), 1015.
DiNardo, J. (2008). Natural experiments and quasi-natural experiments. In S. N. Durlauf, & L. E. Blume (Eds.). *The new Palgrave dictionary of economics*. Basingstoke: Palgrave Macmillan.
Ebrahimi, J., Phan, N., Dou, D., Piniewski, B., & Kil, D. (2016). Characterizing physical activity in a health social network. *Proceedings of the 6th international conference on digital health conference, DH 2016* (pp. 123–129). . https://doi.org/10.1145/2896338.2896349.
Ernala, S. K., Labetoulle, T., Bane, F., Birnbaum, M. L., Rizvi, A. F., Kane, J. M., et al. (2018). Characterizing audience engagement and assessing its impact on social media disclosures of mental illnesses. *Proceedings of the twelfth international conference on web and social media, ICWSM 2018* (pp. 62–71).
Gardner, M. P. (1985). Mood states and consumer behavior: A critical review. *Journal of Consumer Research, 12*(3), 281–300.
Garimella, K., Morales, G. D. F., Gionis, A., & Mathioudakis, M. (2017). The effect of collective attention on controversial debates on social media. *Proceedings of the 2017 ACM on web science conference (WebSci'17)* (pp. 43–52). New York, NY, USA: ACM. https://doi.org/10.1145/3091478.3091486.
Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., & Donaldson, L. (2013). Harnessing the cloud of patient experience: Using social media to detect poor quality healthcare. *BMJ Quality & Safety, 22*(3), 251–255.
Han, X., Wang, L., Crespi, N., Park, S., & Cuevas, Á. (2015). Alike people, alike interests? Inferring interest similarity in online social networks. *Decision Support Systems,*

69(C (January)), 92–106. http://dx.doi.org/10.1016/j.dss.2014.11.008.

Hofman, J. M., Sharma, A., & Watts, D. J. (2017). Prediction and explanation in social systems. *Science, 355*(6324), 486–488.

Khatua, A., Khatua, A., & Cambria, E. (2019). A tale of two epidemics: Contextual word2vec for classifying twitter streams during outbreaks. *Information Processing & Management, 56*(1), 247–257 2019.

Kramer, A. D. I. (2010). An unobtrusive behavioral model of "gross national happiness" *Proceedings of the 28th international conference on human factors in computing systems, CHI 2010* (pp. 287–290). . https://doi.org/10.1145/1753326.1753369.

Kropivnitskaya, Y., Tiampo, K. F., Qin, J., & Bauer, M. A. (2017). The predictive relationship between earthquake intensity and tweets rate for real-time ground-motion estimation. *Seismological Research Letters, 88*(3), 840–850.

Li, C., Lu, Y., Mei, Q., Wang, D., & Pandey, S. (2015). Click-through prediction for advertising in twitter timeline. *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1959–1968). ACM.

Mejova, Y., Haddadi, H., Noulas, A., & Weber, I. (2015). #FoodPorn: Obesity patterns in culinary interactions. *Proceedings of the 5th international conference on digital health 2015 (DH'15)* (pp. 51–58). New York, NY, USA: ACM. https://doi.org/10.1145/2750511.2750524.

Morton, D., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology, 51*(3), 629.

Mueller, J., Jay, C., Harper, S., & Todd, C. (2017). The role of web-based health information in help-seeking behavior prior to a diagnosis of lung Cancer: A mixed-methods study. *JMIR, 19*(6), e189.

Murnane, E. L., & Counts, S. (2014). Unraveling abstinence and relapse: Smoking cessation reflected in social media. *CHI conference on human factors in computing systems, CHI'14* (pp. 1345–1354). . https://doi.org/10.1145/2556288.2557145.

Oktay, H., Taylor, B. J., & Jensen, D. D. (2010). Causal discovery in social media using quasi-experimental designs. *Proceedings of the 3rd workshop on social network mining and analysis, SNAKDD 2009* (pp. 1–9). . https://doi.org/10.1145/1964858.1964859.

Olteanu, A., Varol, O., & Kiciman, E. (2017). Distilling the outcomes of personal experiences: A propensity-scored analysis of social media. *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing, CSCW 2017* (pp. 370–386). . http://dl.acm.org/citation.cfm?id=2998353.

Paul, M. J., & Dredze, M. (2011). You are what you Tweet: Analyzing twitter for public health. *Proceedings of the fifth international conference on weblogs and social media, Barcelona, Catalonia, Spain* July 17–21, 2011.

Paul, M. J., Dredze, M., & Broniatowski, D. (2014). Twitter improves influenza forecasting. *PLoS currents, 6*.

Petty, R. E., Cacioppo, J. T., Sedikides, C., & Strathman, A. J. (1988). Affect and persuasion: A contemporary perspective. *American Behavioral Scientist, 31*(3), 355–371.

Petty, R. E., Priester, J. R., & Wegener, D. T. (1994). Cognitive processes in attitude change. *Handbook of Social Cognition, 2*, 69–142.

Reis, V. L. D., & Culotta, A. (2015). Using matched samples to estimate the effects of exercise on mental health via twitter. *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25–30, 2015* (pp. 182–188). . http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9960.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39*(1), 33–38.

Saha, K., Weber, I., & Choudhury, M. D. (2018). A social media based examination of the effects of counseling recommendations after student deaths on college campuses. *Proceedings of the twelfth international conference on web and social media, ICWSM 2018* (pp. 320–329). . https://aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17855.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics, 25*(1), 1–21.

Stück, D., Hallgrímsson, H. T., Steeg, G. V., Epasto, A., & Foschini, L. (2017). The spread of physical activity through social networks. *Proceedings of the 26th International Conference on World Wide Web, WWW 2017* (pp. 519–528). . https://doi.org/10.1145/3038912.3052688.

Sulistya, A., Sharma, A., & Lo, D. (2016). Spiteful, One-Off, and Kind: Predicting customer feedback behavior on twitter. *International Conference on Social Informatics* (pp. 368–381). Springer.

White, R. W., & Horvitz, E. (2017). Evaluation of the feasibility of screening patients for early signs of lung carcinoma in web search logs. *JAMA Oncology, 3*(3), 398–401.

Zarrinkalam, F., Kahani, M., & Bagheri, E. (2018). Mining user interests over active topics on social networks. *Information Processing & Management, 54*(2), 339–357. https://doi.org/10.1016/j.ipm.2017.12.0034.

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., et al. (2011). Comparing twitter and traditional media using topic models. *In European conference on information retrieval* (pp. 338–349). Springer.