CrossMark

The rippling dynamics of valenced messages in naturalistic youth chat

Seth Frey [1,2,3] · Karsten Donnay [4,5] · Dirk Helbing [5] · Robert W. Sumner [3,5] · Maarten W. Bos [3,6]

Abstract

Even though human behavior is largely driven by real-time feedback from others, this social complexity is underrepresented in psychological theory, largely because it is so difficult to isolate. In this work, we performed a quasi-experimental analysis of hundreds of millions of chat room messages between young people. This allowed us to reconstruct how—and on what timeline—the valence of one message affects the valence of subsequent messages by others. For the highly emotionally valenced chat messages that we focused on, we found that these messages elicited a general increase of 0.1 to 0.4 messages per minute. This influence started 2 s after the original message and continued out to 60 s. Expanding our focus to include feedback loops—the way a speaker's chat comes back to affect him or her—we found that the stimulating effects of these same chat events started rippling back from others 8 s after the original message, to cause an increase in the speaker's chat that persisted for up to 8 min. This feedback accounted for at least 1% of the bulk of chat. Additionally, a message's valence affects its dynamics, with negative events feeding back more slowly and continuing to affect the speaker longer. By reconstructing the second-by-second dynamics of many psychosocial processes in aggregate, we captured the timescales at which they collectively ripple through a social system to drive system-level outcomes.

Keywords

We often think of our effects on one another as discrete transactions: You speak, I respond. A more ecological view sees each person's behavior as triggering successive waves of intra- and interpersonal processes that overlap and interact in complex ways. For example, you may tend to smile after handshakes, partly as an immediate response to feelings of touch, and partly, after a delay, because of the positive interaction that the handshake initiated. You could say that the handshake kicked up more processes in its wake. Waves of processes like these thus operate over various timelines and

sum up to determine your smile from one moment to the next. These unfolding timelines also do not have to come from one source: They can operate over many individuals. From this view, interactions unfold as overlapping ripples in a constantly perturbed pool of psychological dynamics. But what is the timescale of these dynamics? Seconds, minutes, hours, days? And with what strength do they ripple to others and, eventually, back?

Capturing the flow of aggregated dynamics is an essential part of building psychological accounts of social outcomes. Because of their ubiquity and inherently dynamic nature, conversational exchanges are ideal for meeting this challenge (Beckner et al., 2009; Lind, Hall, Breidegard, Balkenius, & Johansson, 2014). Using computational causal inference methods and an archive of months of chat activity from a popular online social game for preadolescents, we isolated the unfolding effects of communication processes as they aggregated and traveled across interacting agents engaged in simple conversations online. Specifically, we computed the time profiles of the influence processes caused by salient chat messages. Our method does not allow us to untangle the bundle of mechanisms that collectively drive the observed dynamics, but through estimates of aggregate effects, it allows us to establish bounds on the strength and duration of any particular mechanism that contributed to the aggregate.

✉ Seth Frey
sfrey@ucdavis.edu

[1] Department of Communication, UC Davis, Davis, CA, USA

[2] Neukom Institute, Dartmouth College, Hanover, NH, USA

[3] Disney Research, Los Angeles, CA, USA

[4] University of Konstanz, Konstanz, Germany

[5] ETH Zurich, Zurich, Switzerland

[6] Carnegie Mellon University, Pittsburgh, PA, USA

In addition to computing chatters' effects on others, we also computed the time profile of a chatter's effects on their self *through* others. We refer to this more time-delayed echoing feedback as *socially mediated self-influence*, because it captures the portion of self-effect that would not have occurred if the original behavior had never affected others. Although many mechanisms of self-influence have been isolated in the laboratory (Cohen & Sherman, 2014; Lind, Hall, Breidegard, Balkenius, & Johansson, 2015; McCann, Higgins, & Fondacaro, 1991; Schimel, Arndt, Banko, & Cook, 2004; Sherman, 2013; Sherman & Cohen, 2006; Van Raalte et al., 1995; Weinberg, Smith, Jackson, & Gould, 1984), isolating one in naturally observed data allows us not only to document the profile of self-influence dynamics, but also to estimate the proportion of all chat that results specifically from one's effect on oneself through others.

Although some have indirectly implied that young people, whose social skills are still developing, might be less sensitive to social influence processes (Simpkins, Schaefer, Price, & Vest, 2013), young people are nonetheless sensitive to their social environment (Krosnick & Judd, 1982; Prinstein & Dodge, 2008). We made several predictions about the dynamics of these influence processes, couched in terms of the time profiles of emotional influence dynamics in the youth-focused online chat system we analyzed. We predicted that influence effects would be evident within seconds of a chat event, and that they would persist—but become weaker—over the scale of multiple minutes. We also expected the same to be true of speakers' indirect feedback effects on themselves upon rippling back through others, but that these effects would be slower and weaker than direct effects. Because affect has clear effects on social influence, we also anticipated an asymmetry in the strength and duration of positively versus negatively valenced chat events.

Literature

The need to capture aggregate psychological dynamics in the real world is urgent. Some of the most fundamental constructs of psychology gain their importance from the tacit or explicit suggestion that a subtle laboratory effect manifests strongly in the wild, as it accumulates and amplifies with repetition and time (Chartrand & Bargh, 1999; Cialdini & Goldstein, 2004; Festinger & Carlsmith, 1959; Hardin & Higgins, 1996; Heatherton, 2011; Leipold et al., 2014; Mason, Conrey, & Smith, 2007; Sherman & Cohen, 2006; Sinclair, Huntsinger, Skorinko, & Hardin, 2005; Wiese, Vallacher, & Strawinska, 2010). At present, the most common strategy for capturing the complexity of an emergent process is to draw a box around the whole system and name a construct after it. Theories of narcissism, self-fulfilling prophecy, reciprocal persuasion, social interdependence theory, self-concept, stereotypes, and cultural

dynamics have all been defined not as individual processes, but as complex systems of many processes (Cialdini, Green, & Rusch, 1992; Davis & Rusbult, 2001; Kashima, 2008; Madon, Jussim, & Eccles, 1997; Markus & Wurf, 1987; Morf & Rhodewalt, 2001; Thompson, Judd, & Park, 2000). Although this "construct-as-system-of-constructs"-based approach has the strength of allowing psychologists to build hierarchies of increasingly abstract constructs, current methods are limited in their ability to capture the richness of the interactions that collectively constitute a given phenomenon in lived experience.

Dynamical theories of psychology

Dynamical psychological theory already provides frameworks for describing phenomena in terms of the temporal trajectories initiated by systems of processes (Beer, 2007; Bingham & Wickelgren, 2008; Gibson, 1966; Vallacher, Nowak, Markus, & Strauss, 1998). In fact, this type of account is essential for modeling certain of the more confounding behaviors that researchers have isolated (Neely, 1977; Weiss, 1953). However, our limited ability to measure the time profiles of social psychological processes continues to hamper the general adoption of dynamical approaches. Specific classes of time profile, such as synchronization processes, have proven relatively simple to study (Hari, Himberg, Nummenmaa, Hämäläinen, & Parkkonen, 2013; Harrison & Richardson, 2009; Lang et al., 2015; Paxton & Dale, 2013; Schmidt, Morr, Fitzpatrick, & Richardson, 2012), as compared to less structured or regular dynamics. A variety of methods continue to offer special insight into psychological dynamics, such as computer simulations (Kashima, Woolcock, & Kashima, 2000; Nowak, Szamrej, & Latané, 1990; Vallacher, Read, & Nowak, 2002), diary and experience sampling methods (Brown & Moskowitz, 1998a; Fraley & Hudson, 2013), naturalistic experiments (Barsade, 2002; Hermans et al., 2012; Vallacher, Nowak, & Kaufman, 1994), field studies (Konvalinka et al., 2011; Xygalatas, Konvalinka, Bulbulia, & Andreas, 2011), and test–retest methods (DiFonzo, Beckstead, Stupak, & Walders, 2016; Huang, Kendrick, & Yu, 2014). Still, these methods remain limited in their ability to explain high-level phenomena—whether these be psychiatric constructs or social outcomes—in terms of basic psychological processes.

Large observational data

Because researchers require additional tools for capturing the time profiles of influence processes in the wild (Brown & Moskowitz, 1998b; Dodds, Harris, Kloumann, Bliss, & Danforth, 2011; Mason et al., 2007; Smaldino, 2014; Vallacher et al., 2002), we proposed to complement the existing toolkit with another approach: causal inference

techniques, as applied to very large naturally occurring observational datasets (Goldstone & Lupyan, 2016), with a special focus on the analysis of online text and social media corpora. The availability of large text corpora of online communication present great opportunities, but are those opportunities to learn about humans, or simply to learn about humans online? Garas, Garcia, Skowron, and Schweitzer (2012) argued for the generalizability of findings in chatrooms beyond their narrow online domain by comparing the use of valenced language on- and offline, and Yee, Bailenson, Urbanek, Chang, and Merget (2007) found evidence that individuals unconsciously maintain their offline gender, interpersonal distance behavior, and eye gaze patterns in online exchanges.

With the present effort, we offer a bridge to the existing body of work in computational social science using large observational network and text datasets to understand the dynamics and micro-to-macro mechanisms of meaning, communication, influence, affect, and language (Lazer et al., 2009).

Online text and language use Within cognitive science, research on language and communication in particular has benefited from the availability of large text datasets and tools for analyzing them. These approaches are succeeding in serving a complex and socially situated vision for the study of language (Beckner et al., 2009).

Doyle and Frank (2015) used the rich data and unique pragmatic constraints of the microblogging platform Twitter to test for a consequence of the linguistic theory of optimal information transmission, work that was extended by Vinson and Dale (2014), who used another constrained social/linguistic context, online restaurant reviews, to show affect-dependent variation in information density. Danescu-Niculescu-Mizil, Lee, Pang, and Kleinberg (2012) and Doyle, Yurovsky, and Frank (2016) used the relatively rigid structure of Twitter conversations to estimate the effects of relative social status and other social factors on linguistic accommodations between interlocutors.

The online possibilities for linguistic scholarship have also attracted more sociological macro-level investigations, whose outcomes are a valuable source of constraints for psychologists, because theories of individual language production and comprehension succeed or fail in explaining population-scale trends. Vilhena et al. (2014) analyzed a large corpus of scientific articles across disciplines to show how jargon reinforces cultural boundaries between academic disciplines, and others have captured macro-scale patterns of semantic evolution over timescales of years to centuries (Hamilton, Leskovec, & Jurafsky, 2016; Hughes, Foti, Krakauer, & Rockmore, 2012). Looking at language change, Gonçalves and Sánchez (2014) used the international popularity of Twitter to provisionally identify new regional dialects of Spanish, whereas Mocanu et al. (2013) examined geographic and seasonal trends of language use variation around the world.

Online text and affect Large text corpora have been valuable for scaling quantitative analyses of affective expression and communication. These developments have been driven by the effectiveness and convenience of tools for automatically extracting affective ratings that correlate sufficiently well with the scores of human raters (Beasley & Mason, 2015; Dodds et al., 2011; Thelwall, Buckley, Paltoglou, Cai, & Kappas, 2010).

Garcia, Kappas, Küster, and Schweitzer (2016) modeled the dynamics of valence and arousal in online discussion forums. Their work builds off of a baseline understanding established by Garas et al. (2012), who showed in a large corpus of chat on Internet relay chat applications that chat rooms have a persistent affective "tone" to which users match their own expressed sentiments, whether positive or negative. Looking at sentiment at a more macro-scale, specifically happiness, researchers at the intersection of computer science and sociology have used Twitter data to estimate population-level happiness trends around the US, along with their correlates with geographic and demographic variables (Mitchell, Frank, Harris, Dodds, & Danforth, 2013), and with word properties and topical events (Dodds et al., 2011).

Online text and influence One of the more active areas of inquiry using large online text data has been the study of social influence, a subject that has attracted interest not just in psychology and cognitive science, but political science, sociology, economics, and information science. Work in this area has represented the cutting edge of research on causal inference in large observational text-driven datasets (Bakshy, Eckles, & Bernstein, 2014; Eckles & Bakshy, 2017), as well as cutting-edge critiques of such quasi-experimental approaches (Cohen-Cole & Fletcher, 2008; Noel & Nyhan, 2011; Shalizi & Thomas, 2011). With its interest in the micro-to-macro manifestations of peer influence and its focus on computational advances in causal inference, the computational social science literature on social influence is very relevant to our present work.

Classic direct social influence has attracted the most attention. Because social influence is of such great interest, and is relatively easy to operationalize and detect, social scientific articles on causal inference often frame methodological advances simultaneously as social influence contributions, using the subject as a cover to establish the validity of a new technique. Aral, Muchnik, and Sundararajan (2009) applied causal inference methods to a large social network dataset in order to disentangle competing mechanisms for the tendency of similar people to associate preferentially, and Lewis, Gonzalez, and Kaufman (2012) approached the same problem using a model-driven approach to the estimation of those alternative explanations. Coviello et al. (2014) creatively used rainfall events as an instrumental variable for estimating emotional contagion over the social network Facebook, a study that some of the same authors

followed with a well-known, very large-scale experiment that manipulated the Facebook "Newsfeeds" of tens of millions of users in order to measure direct emotional influence (Kramer, Guillory, & Hancock, 2014).

In public health, social influence has been proposed to predict changes in smoking, obesity, and happiness (Bliss, Kloumann, Harris, Danforth, & Dodds, 2012; Christakis & Fowler, 2007; Fowler & Christakis, 2008), and studies of user engagement in online support communities have attracted researchers studying consumer behavior, health behavior, and civic engagement (Algesheimer, Dholakia, & Herrmann, 2005; Bakshy, Eckles, Yan, & Rosenn, 2012; Campbell & Kwak, 2010; Gummerus, Liljander, Weman, & Pihlström, 2012; Ruan, Purohit, Fuhry, Parthasarathy, & Sheth, 2012). Timescales of the influence processes documented by these studies tend to be on the order of days and years, with only some work at finer resolution (Garas et al., 2012). Studies over shorter timescales will become more necessary as psychologists take a greater interest in the mechanistic specifics of large-scale influence dynamics.

Moving beyond the well-trod subject of direct social influence are several less direct variations, such as varieties of self-influence, which capture different mechanisms of one's effects on oneself. Recent theories like the "filter bubble" hypothesis—that such social technologies as recommender systems insulate people from challenging ideas—can be interpreted as one's influence on oneself through media technology (Bakshy, Messing, & Adamic, 2015; Liao & Fu, 2013; Pariser, 2011).

Youth As important as these findings are for adult populations, they are even more crucial to understand in the case of young people, who are active Internet users and whose understandings of society and sociality are still developing. Nearly 70% of eight-year-olds in the US go online daily (Gutnick, Robb, Takeuchi, & Kotler, 2010), and minors account for a quarter of the people who live with Internet connections, and possibly for more than a third of cell phone owners (Duggan & Brenner, 2013; File & Ryan, 2014). Whether online behavior predominantly influences real-world behavior, or vice versa, the existence of correlations between on- and offline behavioral patterns are important to monitor and understand (Greenfield & Yan, 2006; Olson, 2010; Prot et al., 2014; Subrahmanyam, Greenfield, Kraut, & Gross, 2001).

The youth-oriented setting of the chat system we examined permitted us to extend our understanding of social influence processes to this underrepresented demographic. Although dynamical perspectives are deeply integrated into developmental research on infants and toddlers (Elman, 2005; Smith, Yu, & Pereira, 2011; Thelen, Schoner, Scheier, & Smith, 2001), similar approaches to older children are less well-developed. The importance of dynamic perspectives has been recognized in work on adolescent social influence

(Brechwald & Prinstein, 2011), but most studies remain either experimental and limited in both scale and naturalism, or large-scale and confined to panel designs whose observations are separated by months or years (Dishion & Tipsord, 2011; Gardner & Steinberg, 2005; Hoeksma, Oosterlaan, & Schipper, 2004; Kandel, 1978; Simpkins et al., 2013; van Geert & Steenbeek, 2005). Pursuing observational phenomena at finer resolutions will make it possible for large-data projects to support and extend laboratory results.

Ethics of using large text data from online social media platforms With the ubiquity of the Internet as a social tool in the lives of even very young people, it is imperative to understand how online engagement interacts with social development. Given the superiority of true random assignment, why would we choose a merely quasi-experimental approach to such effects? The controversy surrounding the Facebook emotion study (Kramer et al., 2014) has raised concerns about inappropriate intrusiveness and the difficulty of eliciting proper informed consent (Kleinsman & Buckley, 2015; Panger, 2016). All researchers who use large online behavioral datasets to study human affairs are complicit, to varying degrees, in today's ethical controversies about privacy online. Large-scale social experiments are on an ethical frontier in this area, and their potential for advancing science is not clearly separable from their potential for abuse. One effect of researchers' reflections on the use of very large social media datasets, with the unprecedented degree of control they offer, has been a reexamination of the need for true experiments in the face of less intrusive alternatives. We follow this trend, and present an alternative to large-scale online experiments: causal inference. Causal inference methods attempt to simulate or reconstruct experimental conditions and can offer mechanistic insight without actually intervening. The "matching" approach we introduce has the additional strength that its results are presented in aggregate in such a way that the final results cannot raise privacy concerns by being deidentified (Narayanan & Shmatikov, 2010). This advantage has been leveraged in other quasi-experimental work with sensitive populations. For example, Schutte and Donnay (2014) used matching on the Wikileaks Iraq War Logs to study the mechanisms that perpetuate violence in conflict zones. In their words,

> We decided that these illegally distributed data could be used in a responsible manner for basic research, given that the empirical analysis would not in any way harm or endanger individuals, institutions, or involved political actors. To ensure this, our analysis only focuses on the events in the statistical aggregate. Moreover, the matching design entails that no marginal effects are estimated for confounding factors, which further strengthens the anonymity of the findings.

Despite these strengths, quality causal inference still depends on the existence of high-granularity behavioral data, and in that sense is still vulnerable to any ethical and privacy concerns that inhere in the mere collection and storage of online activity data.

Data and method

Data

Our data come from a popular proprietary online social game with an international, multilingual user base of millions of young people, most of whom are 8–12 years old. Our analysis is based on about 600,000 events occurring within a corpus of approximately 250 million temporarily archived chat messages from approximately one million players in six language communities (mostly English and Spanish). The data constitute a complete record of many months of activity.

In the game, players are provided with a customizable avatar, a variety of smaller arcade-style games, and a popular chat feature by which they socialize in virtually navigable animated two-dimensional full-screen "rooms." These rooms can host different activities and are linked to each other such that players wander between them in an undirected manner, following their whims or those of others they interact with. The game is more of a "virtual world," in the sense that it does not have a goal and runs continuously at all hours. Players can modify their avatars over time, but they do not gain game experience in any other way, meaning that more experienced players do not have access to a materially larger repertoire of actions. Rooms are usually populated by two to two dozen other players. Chat messages are visible to the other players in each room and visually emanate from each speaker's location, making it easy to follow several overlapping conversations simultaneously. Each message is temporarily stored in a database and tagged with a number of descriptors, many of which are included in Table S1. Due to the game's emphasis on respecting user privacy, it does not permanently archive its data, and it does not solicit or collect any "real-world" individual-level attributes (age, gender, or even country of residence). Because these important features are not available, we controlled for them implicitly by conducting our main analyses within individuals.

Messaging in the game is simple and low bandwidth, with messages limited in length to about 50 characters, and in lexicon to permitted words and combinations of words. To give a sense of the limits of youth-to-youth chat on this system, the 100 most common strings accounted for over 25% of the volume of messages in English, with the top four—"hi," "lol," "ok," and the emoticon "XD"—accounting for 4%. Most chat occurred in the afternoon, with slightly higher activity on weekends. Activity was concentrated on a few servers, and in a few of the many rooms of those servers (Fig. 1). We provide two illustrative sample sequences in Table 1.

Method

After assigning sentiment scores to the corpus, we analyzed it using a causal inference method that is especially suited for aiding causal inference in richly annotated spatiotemporal data. Because the method is relatively new to psychology, we give it substantial attention in this section.

Causal inference with matching

Matching methods rearrange observational data to create experimental conditions artificially (Rubin, 1973). They do this by identifying within the full dataset a small collection of events, assigning these to either "treatment" or "control" categories, and then matching each treatment event to the control that best matches it on every feature besides the binary treatment effect. In this way, treatment and control conditions are built quasi-experimentally from matched data points through a procedure that mimics random assignment. Matching can be seen as approximating random assignment to treatment through a nonlinear, nonrandom correction that counters whatever bias might exist in the properties that distinguish treatment events from controls. Figure 2 illustrates how the conditions of a large-scale laboratory experiment can be reproduced by rearranging data into constructed treatment and control categories. This procedure is passive, in the sense that no experimental manipulation of the environment is performed: Matching operates on populations of events that have already occurred and been logged. Despite its passiveness, attempts to validate matching against experimental results have shown that it performs well with sufficiently large, rich data (Eckles & Bakshy, 2017). Generally, methods such as matching are especially useful in domains in which experimental control is impossible, impractical, or unethical (Schutte & Donnay, 2014).

We implemented matching with several recent refinements—specifically, matched wake analysis with coarsened exact matching—for more automatable, verifiable, scalable, and statistically valid causal inference in a spatiotemporal setting (Blackwell, Iacus, King, & Porro, 2009; Iacus, King, & Porro, 2012, 2015; Schutte & Donnay, 2014).

Social/emotional influence with sentiment analysis of online text chat

We first measured the sentiments of chat statements with the SentiStrength toolkit, an empirically validated sentiment analysis library that has been adapted to many languages (Thelwall, Buckley, & Paltoglou, 2011). SentiStrength returns
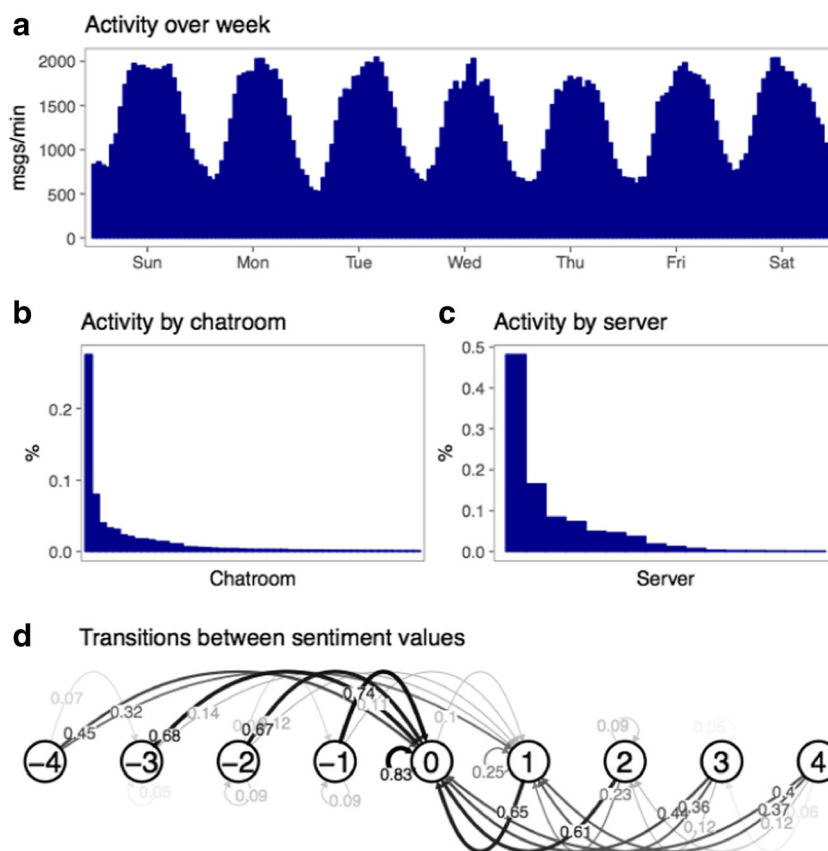
Fig. 1 Descriptive statistics of chat quantity and valence dynamics. Panel A depicts chat quantity aggregated by week. Most activity is during the day. Panels B and C show that chat activity is concentrated into a few chat rooms and servers, with the top ten rooms in a server accounting for more than 50% of its traffic, and the top ten servers accounting for 95% of chat traffic (we omitted several especially low-volume rooms from the tail of panel B). Panel D roughly illustrates the extent of baseline sequential dependence in valenced communication. The nodes from − 4 to 4 are the sums of a message's positive and negative valence scores, and arrows are omitted for transition probabilities less than 5%. As the panel shows, most messages have a net valence of 0, and most valenced messages are followed by a message of valence 0. Sequences of nonzero valence expressions are most common for messages with net valences of 1, − 1, and 2. The Markov chain in this panel illustrates simple dependencies in the sequence of messages, but it does not closely reflect the constructs we used to pose or test our questions

positive and negative sentiment ratings on a 1–5 scale. Higher ratings are increasingly rare, with only 0.34% of messages being assigned the maximum positive or negative rating of 5, and only 0.89% being rated 4 or 5. The library has many options for incorporating sentiment cues from syntactic and semantic features, such as negation and punctuation, but two features stand out that make SentiStrength the most appropriate tool for our corpus. First, it was originally written for Twitter posts, and is therefore appropriate for sentiment analysis in short-form content like the chat messages in our own corpus (which had a 50-character limit). Second, it has been extended to all six of the languages in our corpus. Of the languages in the corpus, English, Spanish, German, and Russian have had some validation, and French and Portuguese have not been validated (Thelwall et al., 2011; Thelwall et al., 2010; Vilares, Thelwall, & Alonso, 2015). For the two languages that have received the most testing against human raters, English and Spanish, the out-of-the-box SentiStrength ratings tend to have 50% correlation with the judgments of human raters, and we assume comparable

correlations for the other languages. See the last panel of Fig. 1 for a rough visual indication of the baseline sequential dependencies in valence from message to message.

A game feature that helps simulate random assignment

The online game we analyzed here is amenable to causal inference because of an automatic chat filter that we call "fake send." Because of the sensitive nature of the population and the ubiquity of inappropriate or unsafe user-generated content, the game's chat feature is engineered to prevent the transmission of such content as insults, cursing, and personally identifying information. When this filter identifies the chat inputs as unacceptable, it notifies users that they have been sanctioned. When this filter identifies inputs as acceptable, it broadcasts them to all other users in the same virtual game room. Finally, when it cannot automatically determine an input to be either acceptable or unacceptable, it applies a "fake send," in which the message seems to the sender to have been sent to others in the room, when in fact it was not. The game's chat monitoring

**Table 1** Sample exchanges with sentiment score and filter status

Exchange	Speaker	Message	Filtered (Fake sent)	Positive sentiment	Negative sentiment
1	A	hi	N	1	− 1
	B	z z z z z z z z z z z z z	N	1	− 1
	A	I'm	N	1	− 1
	C	**ok then _____ great work!**	**Y**	**4**	**− 1**
	B	z z z z	N	1	− 1
	A	I'm the new student	N	1	− 1
	B	z z z z	N	1	− 1
2	A	you left me!	N	2	− 1
	B	piece	Y	1	− 1
	B	OF	N	1	− 1
	C	**lol your lying shes my bff**	**N**	**4**	**− 4**
	D	lmfao	Y	1	− 1
	B	punches	Y	1	− 2
	E	no	N	1	− 1

Illustrative samples of text surrounding two messages (in **bold**) that we defined as events. The first exchange has positive sentiment and was fake sent, for noncompliant syntax. The second was sent and, because positive and negative sentiment are not necessarily opposites, it was scored as both a positive and a negative event. The Speaker column distinguishes speakers in sequence, the Filtered column shows whether a message was excluded from broadcast by the safety filter, which is known to have a high false-positive rate. The positive and negative sentiment columns give the rating of each chat on a 1–5 scale

system is one of many safety features utilized to prevent bullying and age-inappropriate content online. For our purposes, it creates the conditions necessary to isolate and track the effects of single social actions from the continuous stream of chat.

With the very large corpus of past messages that were tagged by the game's chat safety filter, we used matching to approximate random assignment of salient message "events" to treatment ("send") and control ("fake send") conditions. To the extent that the game's safety filter is effective at its job, it is not appropriate to treat it as functioning randomly: It is biased by design. However, by taking advantage of a number of features of the system—including the dimensions along which it is ineffective—and by administering a variety of supplementary robustness checks, we were able to treat the filter as a source of randomizable variation—as variation that can be adjusted to become uncorrelated with other model elements:

- There are technological limits to the effectiveness of the safety filter that make it suitable to treat as a randomizer (Frey, Bos, & Sumner, 2017). Specifically, to guarantee its effectiveness at filtering out unsafe messages, the system's designers have allowed it to be ineffective at passing safe messages; since any failures to filter inappropriate content from a child-focused platform have a high potential cost, to both children and parents, the system conservatively fails to send many statements that a human moderator would not have chosen to filter out. Even though "actually" inappropriate content is very rare, accounting for 1 in 1,000 statements or fewer, most messages

processed by the system (about 60%) are ultimately fake sent. Because the overwhelming majority of fake-sent messages do not differ in content, tone, or appropriateness from sent messages, the overwhelming majority of them are comparable to the sent messages. The matching procedure, described in more detail below, is designed to control for any more subtle bias remaining between them.

- Despite the apparent intrusiveness of the fake-send mechanism, it has not prevented the chat feature from being popular and successful for many years. The reason for this may be that the system has a limited lexicon, and the game's players are young, poor typists, and nonfluent language users with developing social interpretation and interaction skills. Still, some players are likely to notice the silence that followed their fake-sent messages. Our "diff-in-diff" statistical design, described below, controls for such follow-ups to fake-sent messages.

- It is well-established that all users are not equally likely to emit filterable content; certain users are much more likely than others to violate community norms (Cheng, Danescu-Niculescu-Mizil, & Leskovec, 2015). Therefore, we captured the most important source of potential nonrandom differences between conditions (those between individuals) by performing all tests within individuals—a strategy that is possible only because of the large volume of data. This is an important point, since individual-level matching makes it possible to control for the most obvious types of selection bias that our quasi-experimental analysis might introduce.
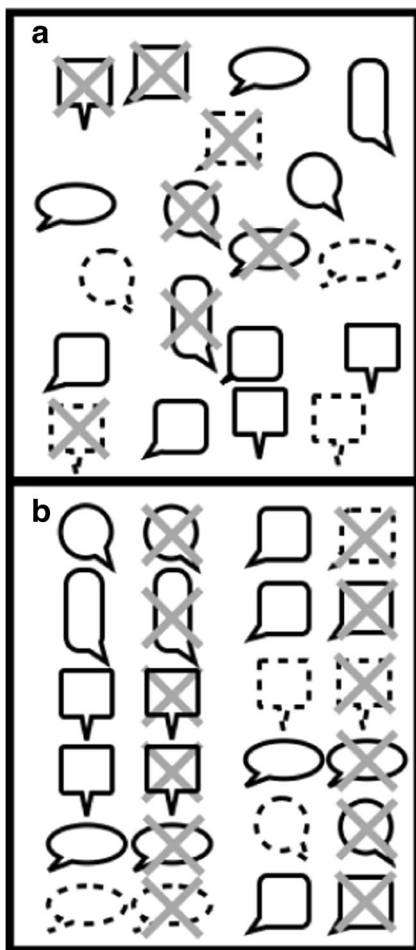
Fig. 2 Matching manages selection bias in observational data to permit causal inference. Panel a depicts an observational dataset of chat events that vary in many attributes. One attribute is that some are sent and some are filtered ("fake sent"). A researcher wants to know the effect of filtering but, for ethical or practical reasons, an experiment is not viable. In panel b, the observations from panel a have been organized such that each sent message is paired with the filtered message that matches it on the most covariates. The result are the paired treatment and control conditions of a quasi-experiment on the effect of filtering: Sent messages are in the treatment and filtered messages are in the control condition

Measuring the wakes of matched messages

Given the chat safety filter as a source of randomizable variation, we proceeded as follows. We first defined within all chat statements a subset of particularly salient chat "events." Events are those messages that received a highly positive or negative sentiment score and were a player's first statement on entering a room. We focused on introductory statements because they are well-separated from each other and less likely to have been caused by prior activity in the room in which they were emitted— more like a ripple caused by an outside pebble than one caused by other ripples. Analyzing the entire corpus, we found 634,000 messages that were a player's first upon entering a room and that had maximum positive or negative valence (490,352 positive events and 143,614 negative

events). We imposed the assumption that strongly emotionally valenced introductory statements would be notable and salient enough to be considered "events" and qualitatively different from chat that was neither introductory nor strongly valenced. Because some of these events were sent and some were fake sent, we could then assign them to treatment or control categories.

We then compared the statistics of many millions of non-event messages before and after events. Our variable of interest, a coarse operationalization of engagement, is the rate of nonevent chat activity in a chat room at any moment, in messages per minute. We measured the rate by counting the number of chats (sent or fake sent) in equal-sized time windows immediately before and after the event. Conversational exchanges in a room are not naturally segmented by the game. The game, and exchanges within it, can continue more or less indefinitely, so our segmentation of conversation in terms of a window around each event was strictly operational. We offer quantity of chat as a simple, if coarse, proxy for user activity that fits within the general scope of social influence, defined broadly as an effect of others on one's behavior, opinions, or emotions. A rate of raw chat activity is far from being a measure with deep psychological meaning, such as personal interest or emotional bonding. However, the quantity of chat is likely to be influenced by a broad array of psychosocial processes, making it well suited to an analysis such as ours, which is interested in the aggregate effects of many processes operating in parallel.

Our analysis is based on a "diff-in-diff" design. Given background messaging rate as the variable of interest, the first step was to calculate a change in the value of the variable before and after each event. This was our dependent variable: the difference in this rate before versus after the event. To conduct a "diff-in-diff" test is to test whether the value of the dependent, a before/after "diff," is itself different between the treatment and control conditions (all model specification details are in the supplemental information [SI]). This type of design is valuable because it controls for the possibility that the "control" condition—sending any kind of message, whether received or not—has an effect of its own. For example, if a player's message is fake sent, and the player notices a conspicuous absence of commentary, the player may perform follow-up actions that cause a different kind of effect on others in the room. By forgoing a test on the change in message rate and testing instead for a difference in that change, we controlled for the unexpected effects of fake-sending a message.

With events, conditions, and dependent variable defined, we can describe the matching procedure. In causal inference by matching, one describes every event in terms of many covariates: their day and time, language, room and server IDs, and other attributes (all described in Table S1). Then, for every treatment event, we identified the control event with the most similar set of covariates. By matching on covariates

in order to build a collection of paired events, this procedure makes it possible to construct quasi-experimental treatment and control conditions from observational datasets.

The motivation behind matching is to make it possible for a regression to permit causal inference without an experimental design. The basic form of a diff-in-diff regression is

$$\text{msgs}_{post} = \beta_0 + \beta_1 \text{msgs}_{pre} + \beta_2 event_{sent} + \in.$$

As is described in Table S1, msgs_{pre} and msgs_{post} give the counts of messages that occurred within some time span before and after the event, and the dummy variable $event_{sent}$ codes whether the salient chat event was treatment or control. The full specification is in the SI ("Model specification").

This analysis tests for the effect of an emotionally salient message on changes in the number of chat statements that followed it, as a function of whether that message was actually sent.

Of course, quasi-experimental studies have been criticized for failing to be experimental studies (Arceneaux, Gerber, & Green, 2010; Miller, 2015; Sekhon, 2009), and their validity is sensitive to the violation of many statistical assumptions, any of which might undermine our analysis. Fortunately, where these violations cannot be prevented, they can at least be diagnosed. Specific simple metrics for diagnosing assumption violations include "overlap rate," "L1 balance," and "local common support" (LCS; Schutte & Donnay, 2014), all of which are elaborated in the "Analysis Details" in the SI. Because of the variety and complexity of these statistics, we used them all and set conservative, easy-to-interpret bounds on the model validity. Specifically, we disqualified any results—regardless of their statistical significance—that reported a total overlap rate (percentage of time windows with multiple treatment or control events) above 10%, or an L1 balance above 0.60, or a common support after matching below 40%. Since such specific bounds can be arbitrary, we conservatively set each bound to be strict relative to what is typically achieved when applying coarsened exact matching to empirical data (e.g., Iacus et al., 2012). The most crucial of these criteria is the overlap rate, which we set particularly low relative to other studies. We also set a conservative significance threshold of $p < .0001$. Although we risked inflating the Type II error rate with our conservative decisions, we think conservatism makes sense, considering the volume of data and the relative novelty of the method for analyzing large datasets in the computational social sciences.

Building time profiles

The above procedure calculates the effect of treatment at a single timescale. To construct a time profile of the unfolding effect of a chat event, we repeated this analysis over multiple timescales. We used multiple time windows of doubling duration, spanning 2 s to 30 min. Specifically, we repeated the analysis by calculating chat messages per minute, before and after an event, within time windows of 2, 4, 8, 16, 30, 60, 120, 240, 480, 960, and 1,920 s. Figures 3 and 4 plot the results of our analyses as repeated over all timescales.

Results

Direct effects

When a player's message is actually sent, as opposed to being fake sent, when does it start affecting others, when does this effect stop, and how strong is it at each moment? In other words, at each timescale, does a strongly positively or negatively valenced chat message cause a subsequent change in the behavior of others, specifically their rate of chat?

For both positive and negative sent events, we found significant increases in the rate of chat by others after a chat event, as compared to when the messages were fake sent. These effects appear in the seconds after an event, and effect sizes remain stable for timescales up to one minute, at which point they start to disappear (Fig. 3). At the 60-s timescale, we found that negative events caused about 0.32 messages per minute (99.9% high-confidence interval [0.24, 0.40]), about three times higher than the effect of positive ones, which cause an increase of 0.091 [0.063, 0.119] messages per minute. Some significant results were disqualified from implying causality. This was mostly because they had too little common support; only at the largest timescales did overlap rates start to drive disqualification.

Figure 3 (and Table S2) shows the results of analyses testing the effects of positive and negative events over several timescales ("Does a sent event with strong [*positive/negative*] valence cause an increase in chat?"). The figure distinguishes between three kinds of results: those that are significant (at $p < .0001$) and respectful of the assumptions of causal inference, those that are nonsignificant, and those technically significant results that we invalidated because they violated causal assumptions.

From these results, we can conclude that the effect of salient introductory chat is to initiate some combination of processes that continue to increase expressed chat over a minute, with the strongest aggregate effects from 2 s after the event.

Socially mediated self-influence

We found that player A's chat influences player B's chat. But if player A's behavior causes an increase in chat by player B, and each elicited chat by player B affects A in turn, and if this process repeats, then it is not sufficient to observe the effect of a single player on others, we must also observe the strengths and timescales of the social feedback processes that
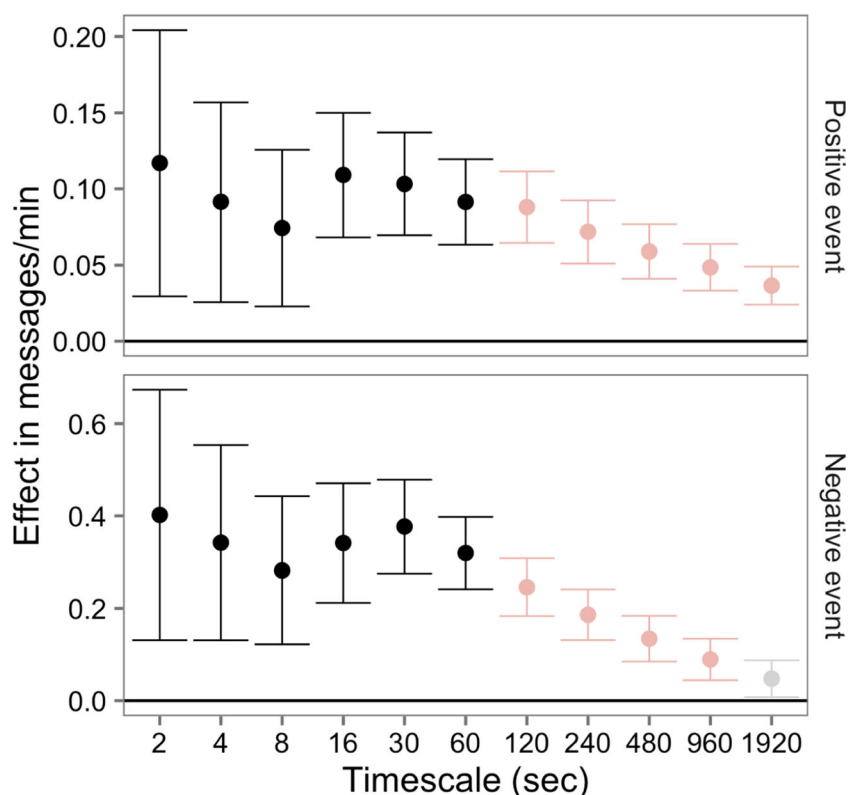
Fig. 3 Time profile of the direct effect of a chat event on subsequent chat activity. Positive events cause an increase of a tenth of a message for up to a minute. Negative events have a much stronger effect, increasing chat by 0.3–0.4 messages per minute. The top panel reports tests across different timescales for the average effect of a very positive chat event on the amount of chat by others in the same chat room. The bottom panel describes the average effect of a very negative statement on subsequent chat by others. The effect of treatment is the average increase in the number of valenced chat statements per minute. Black bars denote analyses that are significant and passed all tests for the validity of causal inference. Gray bars are for analyses that were nonsignificant, and red bars are for analyses that violated the assumptions for causality. Failing tests for causal inference either means that there was not a causal relation or that our method failed to detect one. All statistical tests are significant at $p < .0001$, and bars describe 99.9% confidence intervals. All other figures follow the same format and use the same axes

individual behaviors are driving. Using the fake-send chat mechanism, we isolated one type of feedback process: When a player that is new to a room makes a strongly valenced introductory statement, is the mean amount of that player's own subsequent chat different when the statement was or was not transmitted to others? In other words, when my words affect others, how much of that influence feeds back to me, and when?

Our results support the existence of socially mediated self-influence effects that operate from 8 s to 8 min after a chat event (Fig. 4, Table S3). Although they are specified slightly differently, our tests for direct influence and self-influence were consonant and logically consistent with each other. As compared to the strength of direct influence (Fig. 3; Table S2; SI), the effect sizes and patterns of self-influence were weaker but more extended. Socially mediated self-influence is roughly as strong among negative as among positive events, but they seem to operate on different timescales. Negative events by a player increase that player's subsequent chat on 1- to 10-min timescales (at 60 s, 0.049 messages/min [0.022, 0.77]). The corresponding effect of positive events is 0.017 messages/

min at 60 s [0.005, 0.028] (see Fig. 4), although their effects start to manifest sooner, at 8 s. We interpreted these findings as supporting the existence of socially mediated self-influence among young people.

Effect size and impact

With very large datasets, even trivially small effects may be statistically significant, so it is worth giving special attention to effect size. Therefore, we calculated the proportional as well as the absolute average change in chat after an event. We performed this calculation only for the weaker socially mediated self-influence effect, with the assumption that its relevance would imply the relevance of the stronger direct mechanism that it is based upon. Figure 4 reports the increase in socially mediated self-influence chat after a treatment event, but, regardless of its size, this value could be small relative to the change after a fake-sent control event. At 60 s, which was the most robust point on the analysis timescales, we estimated that the socially mediated self-influence caused by chat events (which are salient but rare) accounts for an average of 1.3% of
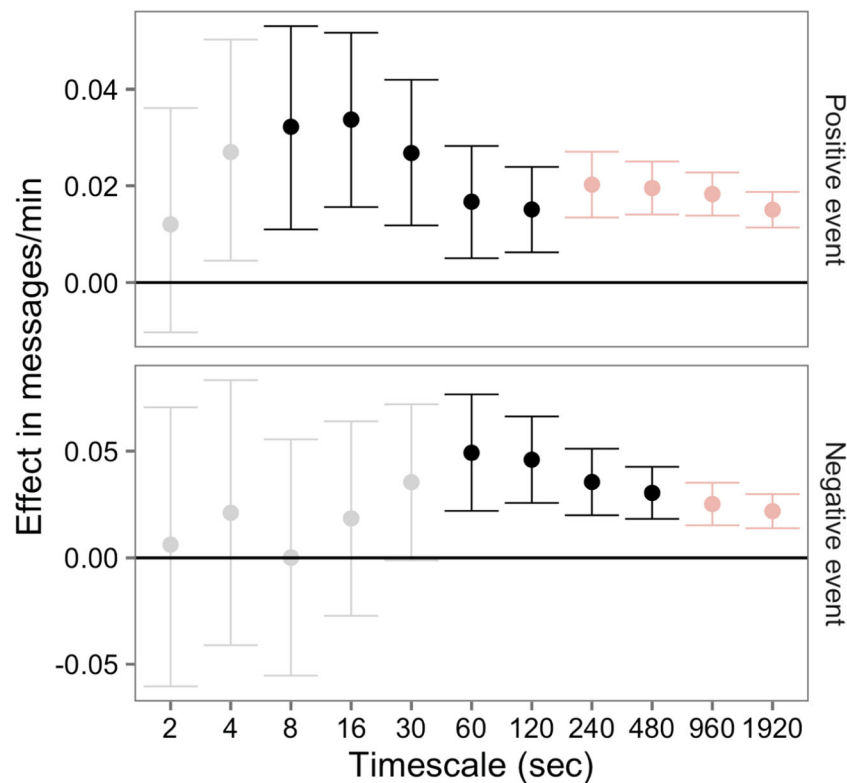
Fig. 4 Time profile of socially mediated self-influence. Between 8 s and 8 min, chat events feed back to affect the original speaker. Negative events are longer-delayed but stronger. The self-effect of both of these types of chat is an increase of between 0.02 and 0.05 messages per minute, corresponding to ~ 1% of the volume of chat. Socially mediated self-influence is the portion of the effect of other people on a speaker that would not have occurred if they had not been affected by that speaker earlier. The top panel reports tests across different timescales for the effect

of a very positive transmitted statement on the number of chat statements by the original speaker. The bottom panel describes the effect of a player's very negative statement on their subsequent moderately negative statements. Analyses coded black are significant and causal, whereas gray analyses are nonsignificant, and red analyses are likely not causal. All tests are significant at $p < .0001$, and intervals describe the area of 99.9% confidence

messages when they are negative, and 0.56% of messages when they are positive. This effect is comparable in size to those that have been experimentally measured in similar domains, such as the Facebook emotion study (Kramer et al., 2014). Presumably, the net effect of socially mediated self-influence is stronger when the full bulk of chat is considered (and not just the chat events around which this analysis was designed).

Robustness

As was specified, several factors remained in the main analysis that might undermine our causal account of the effect of salient chat events on subsequent chat. In general, matching results should not be interpreted without a careful look at whether each analysis satisfied the assumptions of a causal analysis. Since we could not eliminate all possible confounds in one specification, we complement the analysis of the main text with a number of alternative specifications and specialized tests (all reported in the SI). We took

advantage of several recent advances that made it possible to quantify the extent of assumption violations and to eliminate causally invalid analyses (Arceneaux, Gerber, & Green, 2006; Blackwell et al., 2009; Ho et al., 2017; Iacus et al., 2012, 2015; King, Lucas, & Nielsen, 2014; Schutte & Donnay, 2014). Specifically, we tested the ability of our method to establish (1) the robustness of socially mediated self-influence to selection bias due to the confounding selection of low-quality matches, (2) its robustness to selection bias due to nonrandom deletion of data, (3) its robustness to details of the model specification, and (4) its robustness to faulty instruments. These supplemental tests all showed comparable effect sizes over comparable timescales, particularly in the case of negative events.

Discussion

We have reported the timescales over which social influence processes operate in the chat interactions of young

users of an online social game. Using a quasi-experimental design, we found that particularly salient chat events by a focal speaker have a delay of a few seconds before they cause an increase in the chat of others, but that the social and psychological processes initiated by these events continue to drive an increase in the baseline chat rate for about 1 min. And what about the original speaker: What is the feedback effect of this bigger stream of chat messages on the person who elicited it? We found that the feedback takes about 8 s to have a measurable effect, but that it continues for many more minutes. For timescales above 2–8 min (120–480 s in the figures), our results started to become inconclusive, as our time-window-sensitive matching method began to see amplified Type II (false negative) errors. Still, we were able to directly compare the influence dynamics over two to four timescales. Overall, these analyses offer insight into the second-to-second dynamics of chat interactions and how they continually reinitiate themselves over time.

It is most likely that dozens of psychological and social processes are collectively responsible for producing the effects we observed. By measuring their aggregated effect, we put an upper bound on the strength and duration of any single mechanism: no process that had a significant effect on, say, social mediated self-influence behavior, operated more quickly than 8 s, for more than 8 min, or with enough strength to cause more than a 0.4 message per minute increase in chat behavior. In addition to placing an upper bound on the effects on constituent mechanisms of influence ripples, we also propose that our estimates are probably a lower bound on their aggregate strength, that the influence phenomena we isolate can be expected outside of affectively salient introductory chat events. Methodological motives required us to restrict our focus to this very narrow subset of chat, and to entirely ignore the likely possibility that less affectively loaded chats—the overwhelming bulk of interpersonal exchange that we did not define "events" for—are also sending ripples of their own through people's interactions. Some of the robustness checks reported in the supporting material support our findings with looser requirements on the definition of a chat event.

By identifying these complex feedback interactions in the social exchanges of a population that is mostly in the range of 8–12 years old, we also place a lower bound on the level of social complexity that youth in this range can collectively produce. It is especially notable that the rippling dynamics we observe are so strong given that a substantial proportion of the population is still developing a range of relevant skills, from typing to prosocial manners to theory of mind.

Our results advance a view into dynamics that bridges micro-scale psychological processes and lived social experience. Essential to our contribution is the very large data made possible by the Internet, and computational tools that allow us to ethically and rigorously approximate laboratory conditions for real interactions at scale. The computational methods of data science make it possible to make nearly experimental inferences from naturalistic datasets of seemingly intractable complexity (Goldstone & Lupyan, 2016). By documenting the second-to-second dynamics of interaction processes in the "virtual" interactions of real youth, we contribute to a vision for psychological dynamics in which large-data studies of real world settings and controlled laboratory methods complement each other.

Though our results were not derived from a true experiment, causal inference techniques derive from a rigorous theoretical foundation for testing the validity of causal claims. Because they only simulate random assignment, matching methods can offer causal inference without actually intervening in a large-scale social system—a useful property when dealing with vulnerable populations. Matching methods have the additional strength that their results are presented in aggregate in such a way that final results cannot be deidentified. We therefore present our method, which has been applied to other studies of sensitive populations (Schutte & Donnay, 2014), as a more ethical alternative to large-scale online experiments that lack informed consent.

Future work should also be mindful of the methodological limitations of these results. Since quasi-experiments are not experiments, our matching might have had too few features, or lacked important features, and may therefore not have succeeded at a reliable simulation of random assignment. We performed all analyses within-user, implicitly controlling for all possible individual attributes that are stable over a time period of more than an hour, but richer message-level features may reveal new sources of uncontrolled bias. Still, no matter how rich a set of features they use, matching methods will always be susceptible to potential "omitted-variable bias" between sent and fake-sent messages (Shalizi & Thomas, 2011). Also, to get valence scores for our chat data, we applied sentiment ratings that were validated by adults to communications that were generated by children. This does not control for the possibility that statements may elicit systematically different reactions from young people. Youth and adult lexicons do not just differ along the obvious dimension of size, they also differ in the structures of their semantic networks—how the lexicons are internally semantically wired (Beckage, Smith, & Hills, 2011). Hills and others showed evidence in early word-learning lexicons of a preference for words that are well-connected to other known words and to words in the environment (Hills, Maouene, Riordan, & Smith, 2010; Hills, Maouene, Maouene, Sheya, & Smith, 2009). Still, Stella, Beckage, Brede, and De Domenico (2018) found a critical change in the semantic networks of children around the age of 7, younger than most members of our population, after which they more closely resembled those of adults (Steyvers & Tenenbaum, 2005). Reassuringly, the results

among these semantic network studies were comparable despite the fact that some of these studies modeled youth networks by pruning the networks of adults, whereas others calculated youth networks directly. We acknowledge the possibility of important differences between youth and adult chat in this restricted chat system, but we are inclined to invoke the biological basis of emotion and the simple vocabulary of the chat in our corpus in support of an argument that the ratings generated by children and adults should not be so different as to change the direction or strength of our results. Another source of potential concern is that we cannot identify bots in the data, although the child-friendly focus of the platform includes several design decisions that make them rare in the game. The most important limitation of our results is that the raw data and the chat safety filter are proprietary and cannot be shared. As part of the game owner's public commitment to respecting child privacy, all game logs are required to be destroyed a few months after they are generated, meaning that we can only share preprocessed data to those who desire to replicate our results. Likewise, the black-box nature of the safety filter means that we cannot completely rule out the possibility that its decision to fake send a message has some nonrandomness that cannot be controlled for by covariate matching or estimation controls. Still, the filter's very high false-positive rate, combined with its very low miss rate on the rare inappropriate messages, should increase rather than decrease one's confidence in the results we report. Matching methods are useful when random assignment is not available, but the closer the filter is to chance performance, the less bias there is for matching to have to control for.

Conclusion

Online chat data provide extensive archives of granular naturalistic interaction data that we can use to observe how aggregated psychological processes wax and wane from moment to moment. By advancing a multiple-timescale, dynamically extended understanding of psychology, future effort will advance our understanding of how low-level social psychological processes manifest at the scale of lived experience. This work can expand the scope of social-influence-based public health policies and ultimately can help young people respond maturely to social influences, whether positive or negative, online or offline.

References

Algesheimer, R., Dholakia, U. M., & Herrmann, A. (2005). The social influence of brand community: Evidence from European car clubs. *Journal of Marketing*, *69*, 19–34. https://doi.org/10.1509/jmkg.69.3.19.66363

Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, *106*, 21544–21549. https://doi.org/10.1073/pnas.0908800106

Arceneaux, K., Gerber, A. S., & Green, D. P. (2006). Comparing experimental and matching methods using a large-scale voter mobilization experiment. *Political Analysis*, *14*, 37–62.

Arceneaux, K., Gerber, A. S., & Green, D. P. (2010). A cautionary note on the use of matching to estimate causal effects: An empirical example comparing matching estimates to an experimental benchmark. *Sociological Methods and Research*, *39*, 256–282. https://doi.org/10.1177/0049124110378098

Bakshy, E., Eckles, D., & Bernstein, M. S. (2014). Designing and deploying online field experiments. In C.-W. Chung, A. Broder, K. Shim, & T. Suel (Eds.), *Proceedings of the 23rd International Conference on World Wide Web* (pp. 283–292). New York, NY, USA: ACM Press. https://doi.org/10.1145/2566486.2567967

Bakshy, E., Eckles, D., Yan, R., & Rosenn, I. (2012). Social influence in social advertising: Evidence from field experiments. In B. Faltings, K. Leyton-Brown, & P. Ipeirotis (Eds.), *Proceedings of the 13th ACM Conference on Electronic Commerce* (pp. 146–161). New York, NY, USA: ACM Press. https://doi.org/10.1145/2229012.2229027

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*, 1130–1132. https://doi.org/10.1126/science.aaa1160

Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, *47*, 644–675. https://doi.org/10.2307/3094912

Beasley, A., & Mason, W. (2015). Emotional states vs. emotional words in social media. In D. De Roure, P. Burnap, & S. Halford (Eds.), *Proceedings of the ACM Web Science Conference* (Art. No. 31). New York, NY, USA: ACM Press.

Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PLoS ONE*, *6*, e19348. https://doi.org/10.1371/journal.pone.0019348

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., ... Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning, 59*(Suppl.), 1–26.

Beer, R. D. (2007). Dynamical systems and embedded cognition. In K. Frankish & W. M. Ramsey (Eds.), *The Cambridge handbook of artificial intelligence* (pp. 128–148). Cambridge, UK: Cambridge University Press.

Bingham, G. P., & Wickelgren, E. A. (2008). Events and actions as dynamically molded spatiotemporal objects: A critique of the motor theory of biological motion perception. In T. F. Shipley & J. M. Zacks (Eds.), *Understanding events: From perception to action* (pp. 255–285). New York, NY, US: Oxford University Press.

Blackwell, M., Iacus, S. M., King, G., & Porro, G. (2009). cem: Coarsened exact matching in Stata. *Stata Journal*, *9*, 524–546.

Bliss, C. A., Kloumann, I. M., Harris, K. D., Danforth, C. M., & Dodds, P. S. (2012). Twitter reciprocal reply networks exhibit assortativity with respect to happiness. *Journal of Computational Science*, *3*, 388–397. https://doi.org/10.1016/j.jocs.2012.05.001

Brechwald, W. A., & Prinstein, M. J. (2011). Beyond homophily: A decade of advances in understanding peer influence processes. *Journal of Research on Adolescence*, *21*, 166–179. https://doi.org/10.1111/j.1532-7795.2010.00721.x

Brown, K. W., & Moskowitz, D. S. (1998a). Dynamic stability of behavior: The rhythms of our interpersonal lives. *Journal of Personality*, *66*, 105–134. https://doi.org/10.1111/1467-6494.00005

Brown, K. W., & Moskowitz, D. S. (1998b). It's a function of time: A review of the process approach to behavioral medicine research. *Annals of Behavioral Medicine*, *20*, 109–117. https://doi.org/10.1007/BF02884457

Campbell, S. W., & Kwak, N. (2010). Mobile communication and civic life: Linking patterns of use to civic and political engagement. *Journal of Communication*, *60*, 536–555. https://doi.org/10.1111/j.1460-2466.2010.01496.x

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910. https://doi.org/10.1037/0022-3514.76.6.893

Cheng, J., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2015). Antisocial behavior in online discussion communities. In *Ninth International AAAI Conference on Web and Social Media* (pp. 61–70). Retrieved from www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10469

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, *357*, 370–379. https://doi.org/10.1056/NEJMsa066082

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Cialdini, R. B., Green, B. L., & Rusch, A. J. (1992). When tactical pronouncements of change become real change: The case of reciprocal persuasion. *Journal of Personality and Social Psychology*, *63*, 30–40. https://doi.org/10.1037/0022-3514.63.1.30

Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, *65*, 333–371. https://doi.org/10.1146/annurev-psych-010213-115137

Cohen-Cole, E., & Fletcher, J. M. (2008). Detecting implausible social network effects in acne, height, and headaches: Longitudinal analysis. *BMJ*, *337*, a2533. https://doi.org/10.1136/bmj.a2533

Coviello, L., Sohn, Y., Kramer, A. D. I., Marlow, C., Franceschetti, M., Christakis, N. A., & Fowler, J. H. (2014). Detecting emotional contagion in massive social networks. *PLoS ONE*, *9*, e90315. https://doi.org/10.1371/journal.pone.0090315

Danescu-Niculescu-Mizil, C., Lee, L., Pang, B., & Kleinberg, J. (2012). Echoes of power: Language effects and power differences in social interaction. In A. Mille, F. Gandon, & J. Misselis (Eds.), *Proceedings of the 21st International Conference on World Wide Web* (pp. 699–708). New York, NY, USA: ACM Press. https://doi.org/10.1145/2187836.2187931

Davis, J. L., & Rusbult, C. E. (2001). Attitude alignment in close relationships. *Journal of Personality and Social Psychology*, *81*, 65–84. https://doi.org/10.1037/0022-3514.81.1.65

DiFonzo, N., Beckstead, J. W., Stupak, N., & Walders, K. (2016). Validity judgments of rumors heard multiple times: The shape of the truth effect. *Social Influence*, *11*, 22–39. https://doi.org/10.1037/0022-3514.92.5.821

Dishion, T. J., & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychology*, *62*, 189–214. https://doi.org/10.1146/annurev.psych.093008.100412

Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS ONE*, *6*, e26752. https://doi.org/10.1371/journal.pone.0026752

Doyle, G., & Frank, M. C. (2015). Shared common ground influences information density in microblog texts. In *Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL* (pp. 1587–1596). https://doi.org/10.3115/v1/N15-1182

Doyle, G., Yurovsky, D., & Frank, M. C. (2016). A robust framework for estimating linguistic alignment in Twitter conversations. In J. Bourdeau, J. A. Hendler, & R. Nkambou Nkambou (Eds.), *Proceedings of the 21st International Conference on World Wide Web* (pp. 637–648). New York, NY, USA: ACM Press. https://doi.org/10.1145/2872427.2883091

Duggan, M., & Brenner, J. (2013). The demographics of social media users—2012. Pew Research Center.

Eckles, D., & Bakshy, E. (2017). *Bias and high-dimensional adjustment in observational studies of peer effects* (Working article). arXiv: 1706.04692

Elman, J. (2005). Connectionist models of cognitive development: In which next? *Trends in Cognitive Sciences*, *9*, 111–117.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, *58*, 203–210. https://doi.org/10.1037/h0041593

File, T., & Ryan, C. (2014). *Computer and Internet use in the United States: 2013* (No. ACS-28). Washington, DC: US Census Bureau.

Fowler, J. H., Christakis, N. A. (2008). Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ*, *337*, 23–27.

Fraley, R. C., & Hudson, N. W. (2013). Review of intensive longitudinal methods: An introduction to diary and experience sampling research. *Journal of Social Psychology*, *154*, 89–91. https://doi.org/10.1080/00224545.2013.831300

Frey, S., Bos, M. W., & Sumner, R. W. (2017). Can you moderate an unreadable message? "Blind" content moderation via human computation. *Human Computation*, *4*, 78–106. https://doi.org/10.15346/hc.v4i1.5

Garas, A., Garcia, D., Skowron, M., & Schweitzer, F. (2012). Emotional persistence in online chatting communities. *Scientific Reports*, *2*, 402. https://doi.org/10.1038/srep00402

Garcia, D., Kappas, A., Küster, D., & Schweitzer, F. (2016). The dynamics of emotions in online interaction. *Open Science*, *3*, 160059. https://doi.org/10.1098/rsos.160059

Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. *Developmental Psychology*, *41*, 625–635. https://doi.org/10.1037/0012-1649.41.4.625

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.

Goldstone, R. L., & Lupyan, G. (2016). Discovering psychological principles by mining naturally occurring data sets. *Topics in Cognitive Science*, *8*, 548–568. https://doi.org/10.1111/tops.12212

Gonçalves, B., & Sánchez, D. (2014). Crowdsourcing dialect characterization through Twitter. Retrieved April 11, 2018, from www.bgoncalves.com/download/finish/4/108.html

Greenfield, P., & Yan, Z. (2006). Children, adolescents, and the Internet: A new field of inquiry in developmental psychology. *Developmental Psychology*, *42*, 391–394. https://doi.org/10.1037/0012-1649.42.3.391

Gummerus, J., Liljander, V., Weman, E., & Pihlström, M. (2012). Customer engagement in a Facebook brand community. *Management Research Review*, *35*, 857–877. https://doi.org/10.1108/01409171211256578

Gutnick, A. L., Robb, M., Takeuchi, L., & Kotler, J. (2010). *Always connected: The new digital media habits of young children*. New York, NY, USA: Joan Ganz Cooney Center at Sesame Workshop.

Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016, August). *Diachronic word embeddings reveal statistical laws of semantic change*. Article presented at the conference of the Association for Computational Linguistics, Berlin, Germany.

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino & E. T.

Higgins (Eds.), *Handbook of motivation and cognition* (Vol. 3, pp. 28–84). New York, NY: Guilford Press.

Hari, R., Himberg, T., Nummenmaa, L., Hämäläinen, M., & Parkkonen, L. (2013). Synchrony of brains and bodies during implicit interpersonal interaction. *Trends in Cognitive Sciences*, *17*, 105–106. https://doi.org/10.1016/j.tics.2013.01.003

Harrison, S. J., & Richardson, M. J. (2009). Horsing around: Spontaneous four-legged coordination. *Journal of Motor Behavior*, *41*, 519–524. https://doi.org/10.1007/s00221-005-2272-3

Heatherton, T. F. (2011). Neuroscience of self and self-regulation. *Annual Review of Psychology*, *62*, 363–390. https://doi.org/10.1146/annurev.psych.121208.131616

Hermans, R. C. J., Lichtwarck-Aschoff, A., Bevelander, K. E., Herman, C. P., Larsen, J. K., & Engels, R. C. M. E. (2012). Mimicry of food intake: The dynamic interplay between eating companions. *PLoS ONE*, *7*, e31027. https://doi.org/10.1371/journal.pone.0031027

Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language*, *63*, 259–273. https://doi.org/10.1016/j.jml.2010.06.002

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? *Psychological Science*, *20*, 729–739.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2017). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*, 199–236. https://doi.org/10.1162/003465304323023697

Hoeksma, J. B., Oosterlaan, J., & Schipper, E. M. (2004). Emotion regulation and the dynamics of feelings: A conceptual and methodological framework. *Child Development*, *75*, 354–360. https://doi.org/10.1037/0096-1523.12.3.243

Huang, Y., Kendrick, K. M., & Yu, R. (2014). Conformity to the opinions of other people lasts for no more than 3 days. *Psychological Science*, *25*, 1388–1393. https://doi.org/10.1177/0956797614532104

Hughes, J. M., Foti, N. J., Krakauer, D. C., & Rockmore, D. N. (2012). Quantitative patterns of stylistic influence in the evolution of literature. *Proceedings of the National Academy of Sciences*, *109*, 7682–7686. https://doi.org/10.1073/pnas.1115407109

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, *20*, 1–24. https://doi.org/10.1093/pan/mpr013

Iacus, S. M., King, G., & Porro, G. (2015). *A theory of statistical inference for matching methods in applied causal research* (Working paper).

Kandel, D. B. (1978). Homophily, selection, and socialization in adolescent friendships. *American Journal of Sociology*, *84*, 427–436. https://doi.org/10.1086/226792

Kashima, Y. (2008). A social psychology of cultural dynamics: Examining how cultures are formed, maintained, and transformed. *Social and Personality Psychology Compass*, *2*, 107–120. https://doi.org/10.1037/0022-3514.92.2.337

Kashima, Y., Woolcock, J., & Kashima, E. S. (2000). Group impressions as dynamic configurations: The tensor product model of group impression formation and change. *Psychological Review*, *107*, 914–942. https://doi.org/10.1037/0033-295X.107.4.914

King, G., Lucas, C., & Nielsen, R. (2014). The balance-sample size frontier in matching methods for causal inference. *Political Science and Politics*, *42*, 11–22.

Kleinsman, J., & Buckley, S. (2015). Facebook study: A little bit unethical but worth it? *Journal of Bioethical Inquiry*, *12*, 179–182. https://doi.org/10.1007/s11673-015-9621-0

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E.-M., Wallot, S., ... A. Roepstorff (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual.

Proceedings of the National Academy of Sciences, 108, 8514–8519. https://doi.org/10.1073/pnas.1016955108

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences, 111*, 8788–8790. https://doi.org/10.1073/pnas.1320040111

Krosnick, J. A., & Judd, C. M. (1982). Transitions in social influence at adolescence: Who induces cigarette smoking? *Developmental Psychology*, *18*, 359–368.

Lang, M., Shaw, D. J., Reddish, P., Wallot, S., Mitkidis, P., & Xygalatas, D. (2015). Lost in the rhythm: Effects of rhythm on subsequent interpersonal coordination. *Cognitive Science*, *40*, 1797–1815. https://doi.org/10.1111/cogs.12302

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., ... Van Alstyne, M. (2009). Computational social science. *Science, 323,* 721–723. https://doi.org/10.1126/science.1167742

Leipold, B., Bermeitinger, C., Greve, W., Meyer, B., Arnold, M., & Pielniok, M. (2014). Short-term induction of assimilation and accommodation. *Quarterly Journal of Experimental Psychology*, *67*, 2392–2408. https://doi.org/10.1207/s15324834basp0103_2

Lewis, K., Gonzalez, M., & Kaufman, J. (2012). Social selection and peer influence in an online social network. *Proceedings of the National Academy of Sciences*, *109*, 68–72. https://doi.org/10.1073/pnas.1109739109

Liao, Q. V., & Fu, W.-T. (2013). Beyond the filter bubble: Interactive effects of perceived threat and topic involvement on selective exposure to information. In W. E. Mackay (Ed.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2359–2368). New York, NY, USA: ACM Press. https://doi.org/10.1145/2470654.2481326

Lind, A., Hall, L., Breidegard, B., Balkenius, C., & Johansson, P. (2014). Speakers' acceptance of real-time speech exchange indicates that we use auditory feedback to specify the meaning of what we say. *Psychological Science*, *25*, 1198–1205. https://doi.org/10.1177/0956797614529797

Lind, A., Hall, L., Breidegard, B., Balkenius, C., & Johansson, P. (2015). Auditory feedback is used for self-comprehension: When we hear ourselves saying something other than what we said, we believe we said what we hear. *Psychological Science*, *26*, 1978–1980. https://doi.org/10.1177/0956797615599341

Madon, S., Jussim, L., & Eccles, J. (1997). In search of the powerful self-fulfilling prophecy. *Journal of Personality and Social Psychology*, *72*, 791–809. https://doi.org/10.1037/0022-3514.72.4.791

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299–337. https://doi.org/10.1146/annurev.ps.38.020187.001503

Mason, W. A., Conrey, F. R., & Smith, E. R. (2007). Situating social influence processes: Dynamic, multidirectional flows of influence within social networks. *Personality and Social Psychology Review*, *11*, 279–300. https://doi.org/10.1177/1088868307301032

McCann, C. D., Higgins, E. T., & Fondacaro, R. A. (1991). Primacy and recency in communication and self-persuasion: How successive audiences and multiple encodings influence subsequent evaluative judgments. *Social Cognition*, *9*, 47–66. https://doi.org/10.1521/soco.1991.9.1.47

Miller, M. K. (2015). *The uses and abuses of matching* (Working article).

Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The geography of happiness: Connecting Twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE*, *8*, e64417. https://doi.org/10.1371/journal.pone.0064417

Mocanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q., & Vespignani, A. (2013). The Twitter of Babel: Mapping world languages through microblogging platforms. *PLoS ONE*, *8*, e61981. https://doi.org/10.1371/journal.pone.0061981

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry*, *12*, 177–196.

Narayanan, A., & Shmatikov, V. (2010). Myths and fallacies of "personally identifiable information" *Communications of the ACM, 53*, 24–26. https://doi.org/10.1145/1743546.1743558

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254. https://doi.org/10.1037/0096-3445.106.3.226

Noel, H., & Nyhan, B. (2011). The "unfriending" problem: The consequences of homophily in friendship retention for causal estimates of social influence. *Social Networks*, *33*, 211–218. https://doi.org/10.1016/j.socnet.2011.05.003

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review*, *97*, 362–376. https://doi.org/10.1037/0033-295X.97.3.362

Olson, C. K. (2010). Children's motivations for video game play in the context of normal development. *Review of General Psychology*, *14*, 180–187. https://doi.org/10.1037/a0018984

Panger, G. (2016). Reassessing the Facebook experiment: Critical thinking about the validity of Big Data research. *Information, Communication & Society*, *19*, 1108–1126. https://doi.org/10.1080/1369118X.2015.1093525

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. New York, NY: Penguin.

Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal synchrony. *Quarterly Journal of Experimental Psychology*, *66*, 2092–2102. https://doi.org/10.1037/0022-3514.54.6.1063

Prinstein, M. J., & Dodge, K. A. (2008). *Understanding peer influence in children and adolescents*. New York, NY, USA: Guilford Press.

Prot, S., Gentile, D. A., Anderson, C. A., Suzuki, K., Swing, E., Lim, K. M., ... Lam, B. C. P. (2014). Long-term relations among prosocial-media use, empathy, and prosocial behavior. *Psychological Science, 25*, 358–368. https://doi.org/10.1177/0956797613503854

Ruan, Y., Purohit, H., Fuhry, D., Parthasarathy, S., & Sheth, A. P. (2012). Prediction of topic volume on Twitter. In *Proceedings of the International ACM Web Science Conference* (pp. 397–402). New York, NY, USA: ACM Press.

Rubin, D. B. (1973). Matching to remove bias in observational studies. *Biometrics*, *29*, 159–183. https://doi.org/10.2307/2529684

Schimel, J., Arndt, J., Banko, K. M., & Cook, A. (2004). Not all self-affirmations were created equal: The cognitive and social benefits of affirming the intrinsic (versus extrinsic) self. *Social Cognition*, *22*, 75–99. https://doi.org/10.1521/soco.22.1.75.30984

Schmidt, R. C., Morr, S., Fitzpatrick, P., & Richardson, M. J. (2012). Measuring the dynamics of interactional synchrony. *Journal of Nonverbal Behavior*, *36*, 263–279. https://doi.org/10.1007/s10919-012-0138-5

Schutte, S., & Donnay, K. (2014). Matched wake analysis: Finding causal relationships in spatiotemporal event data. *Political Geography*, *41*, 1–10.

Sekhon, J. S. (2009). Opiates for the matches: Matching methods for causal inference. *Annual Review of Political Science*, *12*, 487–508. https://doi.org/10.1146/annurev.polisci.11.060606.135444

Shalizi, C. R., & Thomas, A. C. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods and Research*, *40*, 211–239. https://doi.org/10.1177/0049124111404820

Sherman, D. K. (2013). Self-affirmation: Understanding the effects. *Social and Personality Psychology Compass*, *7*, 834–845. https://doi.org/10.1111/spc3.12072

Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 38, pp. 183–242). Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/S0065-2601(06)38004-5

Simpkins, S. D., Schaefer, D. R., Price, C. D., & Vest, A. E. (2013). Adolescent friendships, BMI, and physical activity: Untangling selection and influence through longitudinal social network analysis. *Journal of Research on Adolescence*, *23*, 537–549.

Sinclair, S., Huntsinger, J., Skorinko, J., & Hardin, C. D. (2005). Social tuning of the self: Consequences for the self-evaluations of stereotype targets. *Journal of Personality and Social Psychology*, *89*, 160–175. https://doi.org/10.1037/0022-3514.89.2.160

Smaldino, P. E. (2014). The cultural evolution of emergent group-level traits. *Behavioral and Brain Sciences*, *37*, 243–254. https://doi.org/10.1017/S0140525X13001544

Smith, L. B., Yu, C., & Pereira, A. F. (2011). Not your mother's view: The dynamics of toddler visual experience. *Developmental Science*, *14*, 9–17. https://doi.org/10.1111/j.1467-7687.2009.00947.x

Stella, M., Beckage, N. M., Brede, M., & De Domenico, M. (2018). Multiplex model of mental lexicon reveals explosive learning in humans. *Scientific Reports*, *8*, 2259. https://doi.org/10.1038/s41598-018-20730-5

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*, 41–78. https://doi.org/10.1207/s15516709cog2901_3

Subrahmanyam, K., Greenfield, P., Kraut, R., & Gross, E. (2001). The impact of computer use on children's and adolescents' development. *Journal of Applied Developmental Psychology*, *22*, 7–30. https://doi.org/10.1016/S0193-3973(00)00063-0

Thelen, E., Schoner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences, 24*, 1–34, disc. 34–86. https://doi.org/10.1017/S0140525X01003910

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, *63*, 163–173. https://doi.org/10.1002/asi.21662

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, *61*, 2544–2558.

Thompson, M. S., Judd, C. M., & Park, B. (2000). The consequences of communicating social stereotypes. *Journal of Experimental Social Psychology*, *36*, 567–599. https://doi.org/10.1006/jesp.1999.1419

Vallacher, R. R., Nowak, A., & Kaufman, J. (1994). Intrinsic dynamics of social judgment. *Journal of Personality and Social Psychology*, *67*, 20–34. https://doi.org/10.1037/0022-3514.67.1.20

Vallacher, R. R., Nowak, A., Markus, J., & Strauss, J. (1998). Dynamics in the coordination of mind and action. In M. Kofta, G. Weary, & G. Sedek (Eds.), *Personal control in action: Cognitive and motivational mechanisms* (pp. 27–59). New York, NY: Springer Science + Business Media. https://doi.org/10.1007/978-1-4757-2901-6_2

Vallacher, R. R., Read, S. J., & Nowak, A. (2002). The dynamical perspective in personality and social psychology. *Personality and Social Psychology Review*, *6*, 264–273. https://doi.org/10.1207/S15327957PSPR0604_01

van Geert, P., & Steenbeek, H. (2005). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review*, *25*, 408–442.

Van Raalte, J. L., Brewer, B. W., Lewis, B. P., Linder, D. E., Wildman, G., & Kozimor, J. (1995). Cork! The effects of positive and negative self-talk on dart throwing performance. *Journal of Sport Behavior*, *18*, 50–57.

Vilares, D., Thelwall, M., & Alonso, M. A. (2015). The megaphone of the people? Spanish SentiStrength for real-time analysis of political tweets. *Journal of Information Science*, *41*, 799–813.

Vilhena, D. A., Foster, J. G., Rosvall, M., West, J. D., Evans, J., & Bergstrom, C. T. (2014). Finding cultural holes. *Sociological Science*, *1*, 221–238.

Vinson, D. W., & Dale, R. (2014). Valence weakly constrains the information density of messages. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 1682–1687). Austin, TX: Cognitive Science Society.

Weinberg, R. S., Smith, J., Jackson, A., & Gould, D. (1984). Effect of association, dissociation and positive self-talk strategies on endurance performance. *Canadian Journal of Applied Sport Sciences*, *9*, 25–32. Retrieved from http://europepmc.org/abstract/med/6705126

Weiss, W. (1953). A "sleeper" effect in opinion change. *Journal of Abnormal and Social Psychology*, *48*, 173–180. https://doi.org/10.1037/h0063200

Wiese, S. L., Vallacher, R. R., & Strawinska, U. (2010). Dynamical social psychology: Complexity and coherence in human experience. *Social and Personality Psychology Compass*, *4*, 1018–1030. https://doi.org/10.1111/j.1751-9004.2010.00319.x

Xygalatas, D., Konvalinka, I., Bulbulia, J., & Andreas, R. (2011). Quantifying collective effervescence: Heart-rate dynamics at a fire-walking ritual. *Communicative and Integrative Biology*, *4*, 735–738. https://doi.org/10.4161/cib.17609

Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & Merget, D. (2007). The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. *Cyberpsychology and Behavior*, *10*, 115–121. https://doi.org/10.1089/cpb.2006.9984