Digital Emotion Contagion

Amit Goldenberg

James J. Gross

Authors Affiliation:

Amit Goldenberg* – Harvard University, Harvard Business School

James J. Gross – Stanford University, Department of Psychology

*Correspondence: agoldenberg@hbs.edu

Key Words: emotion, emotion contagion, digital media, social media

People spend considerable time on digital media, and during this time they are often exposed to others' emotion expressions. This exposure can lead their own emotion expressions to become more like others' emotion expressions, a process we refer to as *digital emotion contagion*. This paper reviews the growing literature on digital emotion contagion. After defining emotion contagion, we suggest that one unique feature of digital emotion contagion is that it is mediated by digital media platforms that are motivated to upregulate users' emotions. We then turn to measurement, and consider the challenges of demonstrating that digital emotion contagion has occurred, and how these challenges have been addressed. Finally, we call for a greater focus on understanding when emotion contagion effects will be strong versus weak or non-existent.

The Ubiquity of Digital Emotion Contagion

In 2014, PNAS published a study that sought to demonstrate emotion contagion on social media using an experimental design [1]. In this study, the content that Facebook users saw was manipulated without their knowledge to be less negative or less positive. Users' emotions were evaluated with a dictionary-based program that counts the number of positive and negative words in each text [2]. Results indicated that those who were exposed to less negative or less positive emotions produced less of these emotions themselves. This is the only published study that has manipulated users' emotions without their knowledge on a digital media platform.

Perhaps fittingly, the emotional response of the general public to this article seemed to illustrate its thesis, as intense emotions become more intense as they spread over social media, bringing more and more users to express their outrage and anxiety about the possibility that their emotions were being manipulated without their explicit consent [3]. The growing outrage expressed by the public eventually led the scientist who authored the report to apologize in a public Facebook post and admit that potential benefits may not have outweighed the costs.

The controversy surrounding this study has drawn increased attention to digital emotional contagion. Growing research in this space highlights the idea that digital emotion contagion occurs in response to a variety of situations, both public and private, and that emotion contagion can play a key role in users' emotions and behavior in a variety of domains. For example, it seems that the digital era we live in has given rise to a large number of online social movements, all highly driven by emotions [4–6], and that emotion contagion is playing a crucial role in driving the spread of these emotions [7]. People also seem to share their personal emotions online in a way that affects not only their own well-being [8], but also the well-being of others who are connected to them [9]. With the tremendous exposure to others' emotions on digital media, the contagious spread of digital emotions seems to play an important role in affecting users' emotions and behavior.

In this paper, we review the growing literature on digital emotion contagion while making two central points. The first point is that digital emotion contagion should be understood as mediated emotion contagion, and that the goals of the digital media companies that serve as its mediators may influence the way digital emotion contagion unfolds in important ways. After providing a general review of emotion contagion in the first section, we dedicate the second section to examining whether and how the mediating role of digital media companies can affect contagion. The second point that we make is that despite its apparent impact on emotional dynamics online, proving that digital emotion contagion has occurred is harder than one might expect. This point in discussed in sections 3 and 4, which focus on the challenges of measuring digital emotion contagion, and the ways these challenges have been met thus far. Finally, we conclude with a section that reviews central findings in this domain, and in light of existing findings, offers new questions and directions for future research on digital emotion contagion.

Emotion Contagion

Emotion contagion has long been recognized as a central driver of individual and collective behavior, as reflected in the writings of philosophers like Hegel [10] and social scientists like Le Bon and Durkhein [11,12]. Within experimental psychology, a seminal book on emotion contagion [13] helped to initiate a wave of empirical research that has sought to specify the nature of emotion contagion and clarify its driving mechanisms [14–16]. Building on previous research, we define emotion contagion as the process by which a perceiver's emotions become more similar to others' emotions as a result of exposure to these emotions. Importantly, we see emotion contagion as a process that can either be conscious or unconscious, with the only necessary condition being that it contributes to increased similarity in emotions between two or more individuals.

Contagion has been shown to occur via at least three mechanisms. The first is mimicry, in which an emotional expression activates synchronous behavior on the part of the perceiver, which in turn activates affective processes [13,17]. Mimicry represents a family of synchronous behaviors that primarily includes facial expressions, but also body postures, eye movements, speech gestures, and laughter [15,18]. The second mechanism is category activation, in which exposure to emotional expressions primes an emotion category, which in turn leads to activation of specific emotional processes [14,19]. Activation is differentiated from mimicry as it does not necessarily involve behaviorally copying an emotional expression, and therefore can result from exposure to emotional cues via other forms of communication such as text. Finally, the third mechanism is social appraisal, in which individuals use others' emotions as a guide for their own emotion appraisals, leading to similar emotional experiences [20,21]. These three mechanisms are not mutually exclusive and can occur in tandem.

Driven by these three mechanisms, emotion contagion can occur as result of many types of exposures to others' emotions. This includes face to face interactions [13], exposure to emotions through text [22,23], and even information gleaned about what other people feel in response to a certain stimulus [24,25]. The variety of mechanisms by which contagion can develop means that it occurs in many different contexts and situations, ranging from interpersonal relationships [26,27] to large collectives [28,29].

For our purposes here, it is useful to limit the scope of emotion contagion and distinguish it from related phenomena. First, emotion contagion may be differentiated from contagion of other, longer term, affective processes such as moods, by focusing on short-term changes in emotions lasting for seconds or minutes [30]. Second, emotion contagion is meant to capture cases in which exposure to other people's emotions leads to similarity in their emotions. This is in contrast to cases in which exposure to others' emotions leads to different or complimentary emotions [31–34], which is especially frequent when individuals are exposed to emotions of people from rival groups [35].

Emotion Contagion on Digital Media

When people interact face-to-face or by phone, their emotional responses are directly perceived by others in an unmediated way. This makes most non-digital interactions different from interactions on digital media, which are almost always mediated by companies who control and manipulate both the content that users see and how they respond to each other. Even on platforms in which there is relatively less management of users' exposure to information, such as online forums, digital news outlets, and video communication platforms, the nature of interactions is guided by top-down design decisions that maximize some behaviors over others. We argue that digital media companies are generally motivated to upregulate users' emotions, and that this potentially amplifies the frequency and intensity of users' exposure to emotions and therefore emotion contagion (Figure 1). These effects may be further amplified by the size and character of digital social networks. Despite exposure to more frequent and more intense emotions, however, it is not yet clear whether and to what degree other processes such as habituation and fatigue act to reduce the strength of digital emotion contagion.

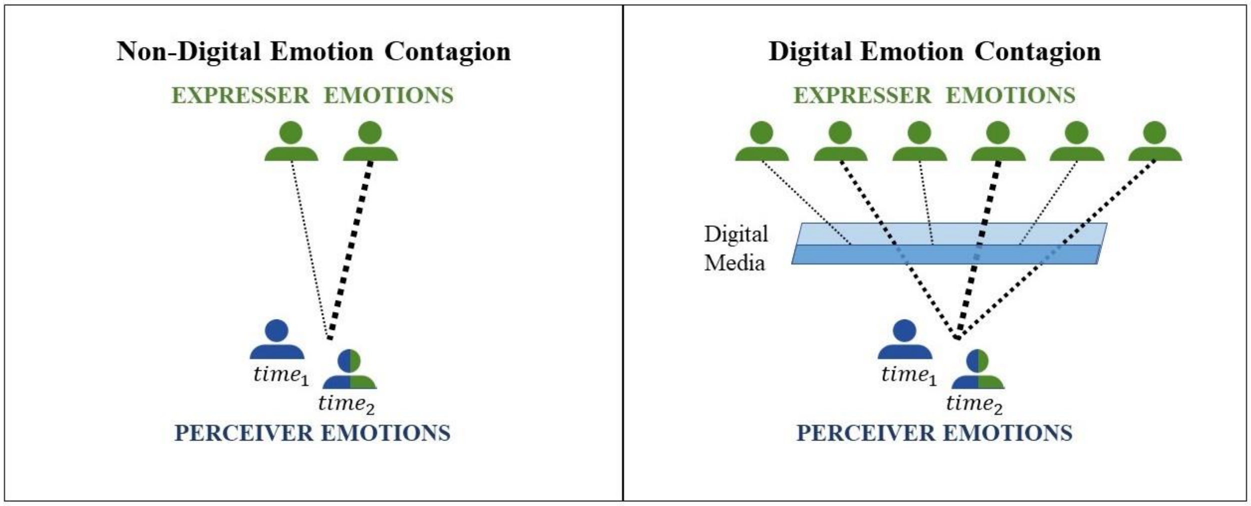

Figure 1. Digital emotion contagion occurs when a perceiver's emotions become more similar to an expresser's emotions over time due to the influence of the expresser's emotions. Digital emotion contagion should be understood as mediated emotion contagion, and the goals of the digital media companies that serve as its mediators lead individuals to be exposed to more intense emotions at a higher frequency.

Figure 2. The number of likes and retweets (log+1 transformed) as a function of the emotions expressed in the tweets (very negative to very positive). We downloaded ~1.5 million random tweets from Twitter API. We then conducted sentiment analysis of the tweets using SentiStrength [112]. The analysis of each text using SentiStrength provides two scores (discrete numbers) ranging from 1-5, one score for positive intensity and one score for negative intensity. We combined the two for -4 (very negative) to 4 (very positive) number. As most tweets do not receive any likes or retweets, we conducted a log+1 transformation to the likes and retweets data. We then fit both a linear and a quadratic functions to the data. Results suggested that the quadratic function was a better predictor of the data, indicating that participants tend to like and retweet emotional tweets compared to non-emotional ones, but also that likes and retweets tend to be higher for positive compared to negative emotions.

Exposure to emotions produced by other users helps to keep users engaged. One of the strongest pieces of evidence for this claim can be seen in the Facebook contagion paper [1], which reports on the *"withdrawal effect,"* in which users have the tendency to produce less content if they are exposed to fewer emotions. If exposing users to others' emotions keeps them engaged, and if engagement is a key outcome for digital media, digital media companies should try to upregulate users' emotions by increasing the frequency and intensity of expressed emotions (particularly positive emotions, see Box 1). This is likely to magnify emotion contagion online.

Increased frequency and intensity of emotion expressions are not only achieved by selectively showing participants more emotional posts, but also by creating an incentive structure that motivates participants to express emotions. Digital media platforms usually incentivize competition for attention and positive reinforcement in the forms of likes or shares [36]. Expressing emotions is an extremely useful way to attract attention and receive likes [4,37–39]. As seen in Figure 2, the intensity of emotional expression predicts the amount of both likes and retweets users receive on Twitter, and this effect is stronger for positive compared to negative emotions (Box 1). The rewards that users receive for expressing emotions create an incentive system that further perpetuates later expression of emotions and therefore contributes to further emotion contagion [40].

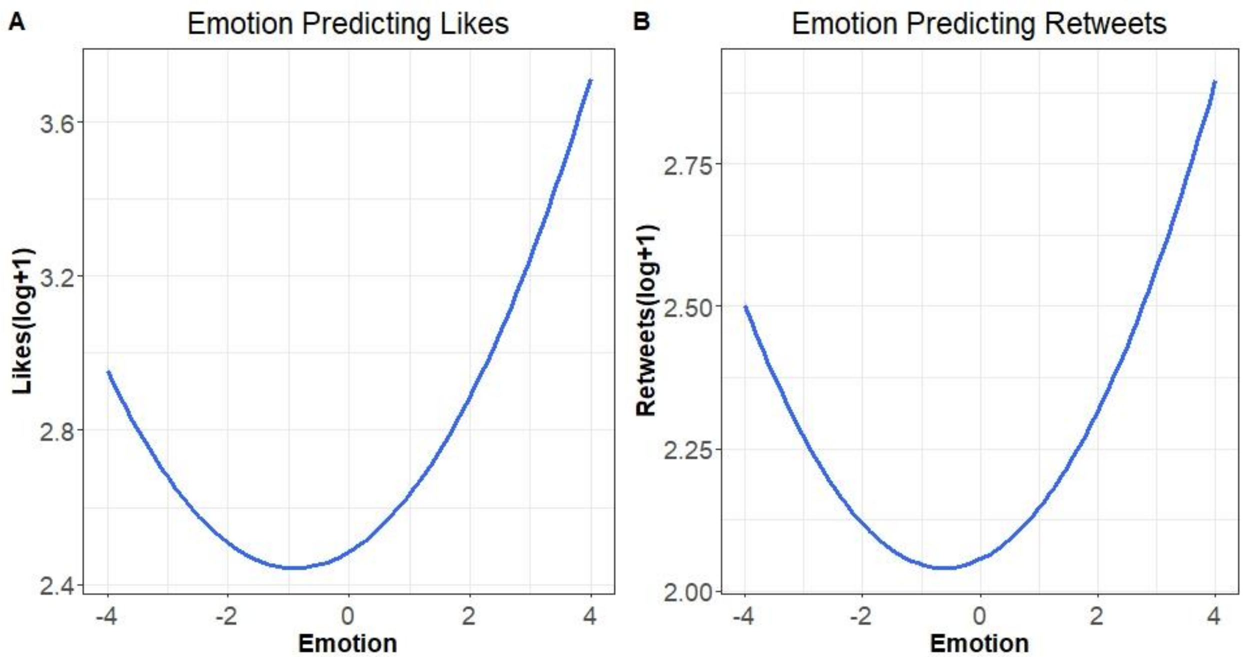

Figure 2. The number of likes and retweets (log+1 transformed) as a function of the emotions expressed in the tweets (very negative to very positive). We downloaded ~1.5 million random tweets from Twitter API. We then conducted sentiment analysis of the tweets using SentiStrength [112]. The analysis of each text using SentiStrength provides two scores (discrete numbers) ranging from 1-5, one score for positive intensity and one score for negative intensity. We combined the two for -4 (very negative) to 4 (very positive) number. As most tweets do not receive any likes or retweets, we conducted a log+1 transformation to the likes and retweets data. We then fit both a linear and a quadratic functions to the data. Results suggested that the quadratic function was a better predictor of the data, indicating that participants tend to like and retweet emotional tweets compared to non-emotional ones, but also that likes and retweets tend to be higher for positive compared to negative emotions.

The overarching goal to increase the frequency and intensity of emotion expressions is combined with the structure of social networks in online platforms in a way that increases digital emotion contagion. Because digital social networks encourage users to connect and interact with as many people as possible, users tend to have larger social networks online and can get exposed to content produced by users who are more distant in their network [41,42]. This often leads to online emotions spreading to a larger number of users and more distant populations [41]. Think for example of the spread of emotions in response to the death of Khaled Mohamad, a 28-year-old Egyptian man from Alexandria who was beaten to death by the Egyptian police in June 2010. Mainstream media did not report the death. However, the Facebook page dedicated to Khaled's death became an occasion for Egyptians to express their frustration to each other. This frustration reached so many people, and increased the intensity of their anger to such a degree, that it led to mass demonstrations in Tahrir Square which contributed to the collapse of the Egyptian government and the initiation of the Arab Spring [43]. As this example suggests, thanks to the mediating role of digital media companies, there is also an increase in the number and the size of social movements which are driven by an ample exchange of emotions, both online and offline [6,36, box 2].

On the face of it, it seems that digital emotion contagion should be more intense, more frequent, and more far-reaching than non-digital contagion. But this may not always be the case. This is because frequent exposure to emotions can also lead to habituation [44] or fatigue [4], making each exposure to emotional expression online less impactful on the perceiver's emotions. Considering that people spend ample amount of time online, they may learn to ignore, at least at some level, the tremendous volume of emotion expressions around them. Furthermore, online social connections tend to be less intimate and valuable to users than offline relationships [45,46], which may also mean that people are less influenced by the emotions of their online friends. So even though digital emotion contagion is likely to be a much more frequent and intense than non-digital emotion contagion, it is also possible that each such exposure to others' emotions on digital media has less of an impact on the individual's emotions. Further work is needed to test whether and under what conditions the activating factors outweigh the inhibiting factors.

Measuring Digital Emotion Contagion: Defining the Challenges

Emotion contagion occurs when a perceiver's emotions become more similar to others' emotions as a result of exposure to these emotions. However, given the number of factors that contribute to changes in a person's emotion from moment to moment [47], and given the typical absence of critical information regarding these other factors when using data from digital media, it is by no means clear when emotion contagion has occurred. It is therefore important both to discuss the challenges that digital emotion contagion researchers face, and to consider the ways in which these challenges have been addressed.

To have confidence that emotion contagion on digital media has occurred, we must be able to successfully address three challenges. The first challenge is to estimate the perceiver's emotion expressions in response to a situation, ideally at two time points, before ($time_1$) and after ($time_2$) the perceiver is exposed to others' emotions. The second challenge is to accurately estimate the perceiver's exposure to an expresser's or expressers' emotions between these two time points. The third challenge is to show that the perceiver's emotion expression was actually influenced by the expresser's emotions as opposed to other sources of influence (such as concurrent changes in the situation itself). Addressing these three challenges can be difficult based on the nature of data available to researchers from social media. In the following section, we discuss each of these challenges in turn, with an eye to how they have been addressed, while recognizing there is no ideal approach and solutions at one level introduce problems at another.

Measuring Digital Emotion Contagion: Meeting the Challenges

When measuring contagion on digital media, the most basic challenge is to estimate the perceivers' emotions based on their digital traces. Importantly, assessing such emotions involves capturing users' expression of emotions, which may be quite different from their emotional experiences, especially on digital media (see Box 3). In the past few years, technology has tremendously improved our ability to estimate emotions by looking at users' facial expressions, vocal responses, and written text (Box 4). The challenge of assessing emotion based on digital traces is compounded by the need to assess a perceiver's emotion at two time points, before and after exposure to an expresser's emotion. In lab experiments, such baseline measurement is relatively easy. However, on digital media it becomes much more challenging, both because finding multiple expressions of emotions by the same user to the same situation is difficult, and because at every time point users are already exposed to some emotions by others. In practice, researchers often ignore perceivers' emotions at $time_1$and measure perceivers' emotions at $time_2$, in light of different emotions expressed by others [[an approach taken also by in-lab experiments, see 25]. For example, in a recent study that examined digital emotion contagion of negative emotions in response to rain, researchers showed that decreases in positive and increases in negative emotions spread to other users who did not experience rain [48], and this was done without examining users' baseline emotional responses. In a recent attempt to establish a pre-exposure measure, researchers estimated users' emotions at $time_1$ by looking at emotions expressed in an earlier content they produced in response to the same situation [7], and this seems to produce stronger estimates of contagion compared to just measuring changes in $time_2$.

Once perceivers' emotions have been estimated, the second challenge is estimating the emotional content that perceivers observed before expressing their own emotions. It is seldom clear what users have encountered. While some users may have been surfing the web for hours, others may have just logged in. Previous studies have taken three different approaches to address this challenge (see Table 1). One approach – referred to as the *window of interest approach* – is to look at the content perceivers could have been exposed to and assume that it was perceived more or less equally across users. For example, in a recent study, researchers [49] estimated the content observed by Twitter users by looking at the average emotional content of tweets produced by a certain perceivers' followees during the hour preceding their posting [7,48,50]. A second approach is the *overall emotional variance approach*. In this approach, researchers focus on macro changes in emotional variance within a certain user community, with the implicit assumption that what every perceiver saw was more or less similar [51,52]. A third approach – referred to as the *emotional cascades approach* – focuses on the people who share or respond to a certain emotional content [53]. In this approach, the assumption is that exposure to emotions elicits similar emotions in perceivers, who then express their emotion by either replying [7,54–57] or further sharing the content [37,58–60]. Emotional cascades resolves the challenge of understanding what a certain perceiver saw (although other content may also influenced their decision), but in the analysis of replies, it is hard to establish that a perceiver replying to an expresser's emotion is reacting to the expresser rather than the situation itself [61]. In the case of shares, it is hard to tell whether sharing an emotional content indicates that the perceiver is feeling similar emotions.

Table 1. Summary of approaches designed to estimate the content observed by the perceiver.

Name	Description	Advantages/Disadvantages	Relevant Papers
Window of interest	Summarizes the emotional content produced by the perceiver's digital community at a certain timeframe prior to the perceiver's expression of emotion at time 2.	Advantages: relatively easy to implement. Disadvantages: we have no clear indication that perceivers actually saw any of the summarized content.	Coviello et al., 2014 Fan, Xu, & Zhao, 2016 Ferrara & Yang, 2015 Goldenberg et al., 2019
Overall variance approach	Measures changes in overall variance of emotions within a	Advantages: provides a view of contagion at the macro level.	Del Vicario et al., 2016 He, Zheng, Zeng, Luo, & Zhang, 2016

	certain digital community over time.	Disadvantages: Other factors may lead to reduction in emotional variance within a community apart from contagion, such as changes in the nature of stimuli or the population within the community.	
Emotional cascade approach	Compares the content produced by replies in relation to their related original post, or counting the amounts of likes and shares produced in response to a certain post.	Advantages: Resolves the challenge of understanding what a certain perceiver saw. Disadvantages: Increases ambiguity about the content that the perceiver produced.	Replies: Chmiel, Sobkowicz, et al., 2011 Dang-xuan & Stieglitz, 2012 Goldenberg et al., 2019 Shares: Alvarez et al., 2015 Brady, Wills, Jost, Tucker, & Van Bavel, 2017 Gruzd, Doiron, & Mai, 2011

Even if one can determine that a perceiver's emotions have changed, and that the perceiver was exposed to the emotions of another user during the period in which the perceiver's emotions changed, one is still left with the third challenge of determining that the emotions of another user played a causal role in that change. One concern is differentiating contagion from similarity-based responses [62]. In a similarity-based response, two or more users are responding to a situation in a similar way not because they are influencing each other, but merely because they are similar to each other.

Perhaps the most compelling way to establish causality is to randomly assign participants to experimental groups that are exposed to the exact same situation, but that differ in their exposure to other users' emotional expressions. This has been done in many in-lab experimental paradigms measuring non-digital emotion contagion [24,25,63]. In field contexts, researchers have tried to address the issue of causal inference in various ways. For example, one recent paper showed that seemingly similar perceivers responded differently to the same situation when exposed to emotions higher or lower in intensity than their own emotions at $time_1$[7]. Other studies have estimated individual and group-level influences within online communities and statistically controlled for the similarity-based effects [64]. However, both methods cannot fully guarantee that we are able to capture contagion [65]. In fact, even the Facebook contagion paper, which manipulated users' perceived content, also struggled with this issue, as manipulating users perceived emotions may also affected the content that they observed [1].

What Predicts the Degree of Emotion Contagion on Digital Media?

Current Findings and Future Directions

Perhaps because measuring the occurrence of emotion contagion on social media is still in its infancy, many studies are still trying to show that contagion exists in a specific platform or situation. In order to move this developing field forward, we believe it will be useful to shift the field's focus toward predicting when emotion contagion will be stronger or weaker. In this section, we summarize what the current literature suggests, and point to gaps in the literature, focusing in turn on the expressed emotion, the network connection, the perceiver, and the platform (see Outstanding Questions).

The strength of emotion contagion is first and foremost dictated by the nature of emotions expressed by the expresser. It is generally assumed that stronger emotion expressions lead to greater emotion contagion. However, there is very little consensus in the literature on what type of emotions lead to stronger contagion. According to the Facebook contagion paper [1], contagion for positive and negative emotions seem to be similar in size, which fits some offline behavioral data [66], and even experiments that examined contagion using neuroimaging [25]. Other findings, however, suggest that positive emotions are more prone to contagion both online [48,49,59,67], as well as offline [68]. These results are somewhat surprising considering the negativity bias, which holds that people tend to pay more attention to negative stimuli [69,70]. We currently know of one study that found that negative emotions, and particularly anger, lead to stronger contagious on digital media [50]. Interestingly, the methods used in this paper were similar to another research project that found stronger contagion for positive emotions [49]. One difference between the two studies is that they measured emotional tweets in different languages, and therefore in different cultural contexts which may differ in their emotion expressions [71]. Based on these conflicting findings, one pressing question is which contexts and cultures lead to more or less emotion contagion for particular, situationally relevant emotions.

The strength of emotion contagion depends not only on the expresser's emotions (intensity and type) but also on the connection between the expresser and the perceiver. It is currently assumed that stronger ties between the expresser and perceiver (evaluated either by reciprocity or by degree of mutual connections) lead to stronger contagion [72]. Yet the relationship between the strength of network connection and contagion seems to depend on the type of expressed emotion [73]. In the first study that tested this question, [50] researchers compared how contagion of anger and joy were influenced by the strength of network connection. They found that anger contagion was stronger in weaker ties compared to joy. Furthermore, a recent study suggests that emotion contagion is not only influenced by network structure but also changes the structure itself [74]. Looking at the spread of negative emotions within an investment company after a drop in stock prices, results suggest that people have a propensity to share their emotions with stronger ties, making these ties even stronger. Future studies that examine the connection between emotion contagion and network structure may be especially important for the advancement of our understanding of the phenomena.

A third critical factor to consider is the perceiver. However, we know little about how perceivers' attributes predict contagion. We therefore wish to suggest a few important future directions. First, the degree of contagion might be influenced by factors such as personality [75,76], which can now be evaluated by users' behavior on social media [77]. For example, it seems likely that people who are more extraverted and agreeable are more likely to catch others' emotions on digital media. It is also likely that users high on neuroticism are more likely to be more influenced by negative emotions in particular [78]. Other individual differences such as status (particularly online status), age, gender, and culture are also likely to influence the degree of contagion between users [51]. Finally, further research should be done on how user characteristics, such as time spent online and degree of activity versus passivity, affect digital emotion contagion. For example, a recent study examining emotions in online communities suggests that more active users tend to shift more quickly to express negative emotions [52]. Future work should further examine these questions.

Finally, the type of platforms that users employ, each with its slightly different set of motivations in relation to a desired level of users' emotion, and the type of content they produce in these platforms is also likely to influence the nature of contagion[79,80]. Different digital media platforms are characterized by different emotional baselines (Box 1), which may affect the degree of contagion of certain emotions. Social media platforms and video sharing sites like YouTube are often characterized by more positive emotions [1,48,64,81], although this depends on the specific content [82]. Online forums tend to be more positive as well, but forums that are centered around well-being, depression, and anxiety are more likely to be negative, primarily due to the emotional baseline of the users that create the content [83]. Comments in responses to online newspaper articles tend to include a larger mix of emotions, and some of them tend to be negative [84] while others are more positive [67]. The emotional content of the situations that are common in digital spaces can play a role in emotion contagion. If negative situations are present in the vast majority of situations, users are more likely to be influenced by more negative emotions [7]. However, remember that although the Facebook contagion study reported much stronger positive emotions, no differences in contagion effect sizes were found, suggesting the more research should be done to answer these questions [1].

Concluding Remarks

The goal of the current project was to review the growing literature on digital emotion contagion while making two central points. The first point is that digital emotion contagion should be understood as mediated emotion contagion. The goals of digital media companies – to increase the frequency and intensity of users' emotions – are likely to act as excitatory factors for digital emotion contagion. However, increased exposure may also contribute to habituation and fatigue, especially considering the fact that social connection on digital media are less meaningful, and therefore inhibit digital emotion contagion. Future work should examine these different features of digital emotion contagion and their impact on the degree of contagion (see Outstanding Questions).

The second point that we make is that despite its apparent impact on emotional dynamics online, proving that digital emotion contagion has occurred is harder than one might expect. For example, users can have similar emotional responses to similar situations without any contagion, but differentiating such cases of similar emotional responses from contagion is extremely challenging. It is therefore important to measure contagion in different ways, while recognizing the advantages and disadvantages of any measurement.

It is likely that because proving that digital contagion actually occurred is challenging, most existing studies try to prove that contagion actually occurred. We believe that with a few established methods, it is now time to shift the field's focus toward predicting when emotion contagion will be stronger or weaker. Future studies should ask what type of expressed emotions, expressed by whom, to whom, and in what platform can predict stronger or weaker contagion. We are excited by the opportunities ahead in this growing field supported by an ever increasing data and use of digital media.

Box 1: Positive Emotions on Digital Media

People express all manner of emotions on social media, but overall, they seem to express more positive emotions than negative emotions. For example, in the Facebook contagion paper, 46.8% of all analyzed posts contained positive words and 22.4% contained negative words [1], and the same is true for other digital media platforms [64,81,85]. The tendency to express more positive emotions is thought to arise both as a result of users' internal motivations, and as a result of top-down regulation by digital media platforms.

In general, people tend to prefer to feel and therefore express positive emotions because it feels better and because in most social interactions, expressing positive emotions is more helpful in advancing the individual social goals both online and offline [86]. Congruent with this idea is the finding that expression of positive emotions is generally perceived as more appropriate than the expression of negative emotions [81]. For example, a recent study that showed that users perceived the positive emotions of joy and pride to be the most appropriate emotions on Facebook, Twitter, Instagram, and WhatsApp, and perceived the negative emotions of sadness and anger to be the least appropriate emotions. Second, digital media interactions are often driven by social comparison [87], and expressing positive emotions proves one's success and therefore helps the individual to positively compare themselves to others.

In addition to users' internal motivation to express positive emotions, digital media companies are also motivated to increase users' positive emotions. Users produce more content when they are exposed to positive versus negative emotions [1]. Therefore, the design of many digital media platforms contributes to a positive bias in emotion expression. In most social media platforms, participants can express their enjoyment or gratitude in response to content by liking it

[88], but there are no “unlike” buttons, which leads to more positive than negative feedback. In addition to these explicit design features, and with the evidence that digital media companies wish to maximize users’ emotions, some suggest that digital media algorithms particularly promote content with positive emotions. For example, some have suggested that Facebook chose to bury posts from the Ferguson unrest, a social movement that grew after the death of Michael Brown in Ferguson, Missouri, in favor of more positive posts [6]. However, providing empirical evidence for such claims is extremely challenging, as digital media algorithms remain a black box for external researchers.

Box 2: Digital Media, Emotion Contagion, and Social Movements

Increased use of digital media, especially social media, has transformed the way social movements unfold. In particular, users’ large number of social connections and high frequency of social interactions are leading to more frequent online social movements [6,89]. It is almost impossible to imagine movements such as the Arab Spring or the Black Lives Matter without digital media.

One important driver for social movements is the exchange of emotions between users, particularly anger [4,5,90]. Anger tends to spread faster than other emotions on social media [50] and to cascade to more users by shares and retweets, enabling quicker distribution to a larger audience [37,58]. Users are also motivated to share their anger because they wish to signal their social network about their morality [4,7] and to convince them to join the movement [31].

Despite the obvious impact that digital media has on online social movements, the translation of online activity to collective action outside social media is often surprisingly limited. For example, while the Save Darfur Facebook campaign – designed to increase awareness and donations to the war in Sudan – was able to recruit 1.2 million members to the movement, the amount and quality of activism that resulted from the campaign were relatively modest as most users did not donate money for the cause [91]. A similar example can be seen in the viral Ice Bucket Challenge, which was designed to raise awareness and donations to ALS. Over 28 million users joined the challenge, and $115 million were raised. Yet, donations were not sustained fell back to pre-campaign levels the year after [38]. Furthermore, 1 out of 4 users who completed the challenge did not mention ASL in their videos and only 1 out of 5 mentioned a donation, suggesting that much of the public interest was not translated into actual action. These examples show that while it is clear that digital media greatly contributes to online social movements, the question of how much these movements translate to action in the real world is still open.

Box 3: Emotion Experience versus Expression on Digital Media

To what extent do people's emotion expressions on digital media reflect their true emotional experiences? After all, we know that emotion experience and expression is imperfectly correlated in everyday life [92], and there are factors that might either increase or decrease the gap between experience and expression in digital versus non-digital contexts.

The prevailing assumption is that communication on digital media allows more opportunities for positive self-presentation than in non-digital contexts [93–95]. If so, we might assume that expression-experience differences are larger online, such that online users either upregulate or downregulate emotional expressions to fit with self-presentational goals [96]. This can be supported by a few arguments. First, communication on digital media is asynchronous, such that users do not have to respond to each other in real time [97,98]. Longer response time provides more opportunity for expression regulation, which may increase the difference between experience and expression. Second, digital communication also involves a larger audience [95], which often leads to an increase in self-presentation motivations [99]. Finally, digital communication allows for less information richness [80,97,100], which means that users have to amplify their expressions to make sure that perceivers understand their emotions.

On the other hand, digital media also provide opportunities for self-disclosure and genuine expression of emotions in ways that are hard to come by in non-digital contexts [95,97,101–103]. First, digital media allows people to express themselves in anonymity, which seems to promote self-disclosure [103–105]. Second, online users can receive a much a larger amount social support from their social environment, particularly in cases in which their offline social environment does not support these emotions. This can sometimes lead to upregulation of emotion expression with the goal of getting more likes [4,40], but can also lead to genuine self-disclosure in ways that could not occur in face-to-face interactions [101]. Finally, some argue that due to accountability and feedback provided by users' social networks, digital media represents an extension of users' real life and that people communicate their true self [80,106,107]. These considerations suggest that experience-expression differences in digital contexts may be either smaller than or similar to face-to-face interactions.

Box 4: Estimating Users' Emotions from Digital Media

Digital media activity allows researchers to detect users' emotions from different types of signals [108]. Facial expressions, produced in videos and photos, can be analyzed by image sentiment analysis software [109]. Audio can be analyzed in terms of pitch [110]. And text produced by users can be analyzed used text-based sentiment analysis tools [111].

Of these response channels, text is the most commonly analyzed. For this reason, text-based sentiment analysis tools have received the most attention. These tools vary greatly in the way they process text. Some sentiment analysis tools, such LIWC, count emotional words based on pre-determined dictionaries [2], others add certain context rules for such dictionaries and include more complicated word compositions [112,113], and some use sophisticated machine learning algorithms [114]. The comparison between these tools is challenging as outcomes may depend on the specific domain (product review, social media posts etc.), and the length of the text (twitter posts versus blogs) [115]. It is likely that machine learning algorithms that are trained to predict emotions in product reviews would be superior to other tools at predicting emotions in their pre-trained domain, but without fine tuning, these algorithms may prove to be inferior to more basic tools in predicting emotions in a completely different domain. In general, however, there are many tools with good predictive power that correlate well with raters' emotion ratings. The sentiment analysis tool VADER, for example, achieved a correlation of r 0.88 with human raters in the rating of 4,000 tweets in the task of classifying tweets into positive, negative or neutral.

One especially important component of emotions expressed in text is the use of emotion icons (emoticons, emojis), which are visual representations of various emotional states (and other states). Emoticons are extremely popular and used by 92% of online population [116]. The use of emoticons is extremely helpful from a communicative perspective as it provides a relatively clear picture of the emotions that participants wish to express [97,117]. Emoticons are also perceived in a similar way to facial expressions [118]. While some basic sentiment analysis tools do not incorporate emoticons in their analysis, most newer tools take them into account and this improves their ability to estimate users emotions from their text [113].

Acknowledgment: The authors wish to thank Yuan Chang Leong, Aharon Levy, Tamar Saguy and David Garcia for their constructive feedback during the preparation of this manuscript.

References

1 Kramer, A.D.I. *et al.* (2014) Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci.* 111, 8788–8790

2 Pennebaker, J.W. *et al.* (2015) *The development and psychometric properties of LIWC2015*,

3 Panger, G. (2016) Reassessing the Facebook experiment: critical thinking about the validity of big data research. *Inf. Commun. Soc.* 19, 1108–1126

4 Crockett, M.J. (2017) Moral outrage in the digital age. *Nat. Hum. Behav.* 1, 769–771

5 van Zomeren, M. *et al.* (2012) Protesters as "passionate economists": A dynamic dual pathway model of approach coping with collective disadvantage. *Personal. Soc. Psychol. Rev.* 16, 180–199

6 Tufekci, Z. (2017) *Twitter and tear gas : the power and fragility of networked protest*, Yale University Press.

7 Goldenberg, A. *et al.* (2019) Beyond emotional similarity: The role of situation specific

motives. *J. Exp. Psychol. Gen.*

8 Lomanowska, A.M. and Guitton, M.J. (2016) Online intimacy and well-being in the digital age. *Internet Interv.* 4, 138–144

9 Hill, A.L. *et al.* (2010) Emotions as infectious diseases in a large social network: the SISA model. *Proc. R. Soc. London B Biol. Sci.* 277,

10 Taylor, C. (1975) *Hegel*, Cambridge University Press.

11 Le Bon, G. (1896) *The crowd: A study of the popular mind*, Viking.

12 Durkheim, É. (1912) *The elementary forms of religious life,* The Free Press.

13 Hatfield, E. *et al.* (1994) *Emotional contagion*, Cambridge University Press.

14 Peters, K. and Kashima, Y. (2015) A multimodal theory of affect diffusion. *Psychol. Bull.* 141, 966–992

15 Parkinson, B. (2011) Interpersonal emotion transfer: Contagion and social appraisal. *Soc. Personal. Psychol. Compass* 5, 428–439

16 Barsade, S.G. *et al.* (2018) Emotional contagion in organizational life. *Res. Organ. Behav.* 38, 137–151

17 Hess, U. and Fischer, A.H. (2014) Emotional mimicry: Why and when we mimic emotions. *Soc. Personal. Psychol. Compass* 8, 45–57

18 Prochazkova, E. and Kret, M.E. (2017) Connecting minds and sharing emotions through mimicry: A neurocognitive model of emotional contagion. *Neurosci. Biobehav. Rev.* 80, 99–114

19 Niedenthal, P.M. *et al.* (2009) Embodiment of emotion concepts. *J. Pers. Soc. Psychol.*

96, 1120–1136

20 Manstead, A. and Fischer, A.H. (2001) Social appraisal: The social world as object of and influence on appraisal processes. In *Emotion: Theory, methods, research* (Scherer, K. R. et al., eds), pp. 221–232, Oxford University Press

21 Clément, F. and Dukes, D. (2017) Social appraisal and social referencing: Two components of affective social learning. *Emot. Rev.* 9, 253–261

22 Cheshin, A. *et al.* (2011) Anger and happiness in virtual teams: Emotional influences of text and behavior on others' affect in the absence of non-verbal cues. *Organ. Behav. Hum. Decis. Process.* 116, 2–16

23 Guillory, J. *et al.* (2011) Upset now?: Emotion contagion in distributed groups. *Proc. SIGCHI Conf. Hum. factors Comput. Syst.* DOI: 10.1145/1978942.1979049

24 Lin, L.C. *et al.* (2018) Intergroup social influence on emotion processing in the brain. *Proc. Natl. Acad. Sci.* 115, 10630–10635

25 Willroth, E.C. *et al.* (2017) Social information influences emotional experience and late positive potential response to affective pictures. *Emotion* 17, 572–576

26 Feldman, R. and Klein, P.S. (2003) Toddlers' self-regulated compliance to mothers, caregivers, and fathers: Implications for theories of socialization. *Dev. Psychol.* 39, 680–692

27 Beckes, L. and Coan, J.A. (2011) Social baseline theory: The role of social proximity in emotion and economy of action. *Soc. Personal. Psychol. Compass* 5, 976–988

28 Páez, D. *et al.* (2015) Psychosocial effects of perceived emotional synchrony in collective

gatherings. *J. Pers. Soc. Psychol.* 108, 711–729

29 Konvalinka, I. *et al.* (2011) Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proc. Natl. Acad. Sci.* 108, 8514–8519

30 Fowler, J.H. and Christakis, N.A. (2009) Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ* 338, 23–26

31 Goldenberg, A. *et al.* (2014) How group-based emotions are shaped by collective emotions: Evidence for emotional transfer and emotional burden. *J. Pers. Soc. Psychol.* 107, 581–596

32 Alshamsi, A. *et al.* (2015) Beyond contagion: Reality mining reveals complex patterns of social influence. *PLoS One* 10, 1–15

33 Iyengar, S. *et al.* (2018) The origins and consequences of affective polarization in the United States. *Annu. Rev. Polit. Sci.* 22, 1–35

34 Li, X. and Hitt, L.M. (2008) Self-selection and information role of online product reviews. *Inf. Syst. Res.* 19, 456–474

35 Cikara, M. and Fiske, S.T. (2013) Their pain, our pleasure: Stereotype content and schadenfreude. *Ann. N. Y. Acad. Sci.* 1299, 52–9

36 Tufekci, Z. (2013) “Not this one”: Social movements, the attention economy, and microcelebrity networked activism. *Am. Behav. Sci.* 57, 848–870

37 Alvarez, R. *et al.* (2015) Sentiment cascades in the 15M movement. *EPJ Data Sci.* 4, 1–13

38 van der Linden, S. (2017) The nature of viral altruism and how to make it stick. *Nat. Hum.*

Behav. 1, 1–3

39 Brady, W.J. *et al.* (2019) Attentional capture helps explain why moral and emotional content go viral. *J. Exp. Psychol. Gen.* DOI: 10.1037/xge0000673

40 Brady, W.J. *et al.* (2019) The MAD model of moral contagion: The role of motivation, attention and design in the spread of moralized content online. DOI: 10.31234/OSF.IO/PZ9G6

41 Watts, D.J. (2004) *Six degrees: The science of a connected age*, W.W. Norton & Company.

42 Ahn, Y.Y. *et al.* (2007) , Analysis of topological characteristics of huge online social networking services. , in *16th International World Wide Web Conference, WWW2007*, pp. 835–844

43 Ghonim, W. (2012) *Revolution 2.0: The power of the people is greater than the people in power*, Houghton Mifflin Harcourt Publishing Company.

44 Wilson, T.D. and Gilbert, D.T. (2008) Explaining Away: A model of affective adaptation. *Psychol. Sci.* 3, 370–386

45 Cummings, J.N. *et al.* (2002) The quality of online social relationships. *Commun. ACM* 45, 103–108

46 Schiffrin, H. *et al.* (2010) The associations among computer-mediated communication, relationships, and well-being. *Cyberpsychology, Behav. Soc. Netw.* 13, 299–306

47 Kuppens, P. *et al.* (2010) Feelings change: Accounting for individual differences in the temporal dynamics of affect. *J. Pers. Soc. Psychol.* 99, 1042–1060

48 Coviello, L. *et al*. (2014) Detecting emotional contagion in massive social networks. *PLoS One* 9, e90315

49 Ferrara, E. and Yang, Z. (2015) Measuring emotional contagion in social media. *PLoS One* 10, e0142390

50 Fan, R. *et al.* (2016) Higher contagion and weaker ties mean anger spreads faster than joy in social media. at <http://arxiv.org/abs/1608.03656>

51 He, S. *et al.* (2016) Exploring entrainment patterns of human emotion in social media. *PLoS One* 11, e0150630

52 Del Vicario, M. *et al.* (2016) Echo chambers: Emotional contagion and group polarization on Facebook. *Sci. Rep.* 6, 1–12

53 Rimé, B. *et al.* (1998) Social sharing of emotion: New evidence and new questions. *Eur. Rev. Soc. Psychol.* 9, 145–189

54 Chmiel, A. *et al.* (2011) Negative emotions boost user activity at BBC forum. *Phys. A Stat. Mech. its Appl.* 390, 2936–2944

55 Dang-xuan, L. and Stieglitz, S. (2012) Impact and diffusion of sentiment in political communication – An empirical analysis of political weblogs. *Proc. Sixth Int. AAAI Conf. Weblogs Soc. Media*

56 Friedman, R. *et al.* (2004) The positive and negative effects of anger on dispute resolution: Evidence from electronically mediated disputes. *J. Appl. Psychol.* 89, 369–376

57 Van Kleef, G.A. *et al.* (2004) The interpersonal effects of anger and happiness in negotiations. *J. Pers. Soc. Psychol.* 86, 57–76

58 Brady, W.J. *et al.* (2017) Emotion shapes the diffusion of moralized content in social networks. *Proc. Natl. Acad. Sci.* 114, 7313–7318

59 Gruzd, A. *et al.* (2011) Is happiness contagious online? A case of twitter and the 2010 Winter Olympics. *Proc. Annu. Hawaii Int. Conf. Syst. Sci.* DOI: 10.1109/HICSS.2011.259

60 Christophe, V. and Rimé, B. (1997) Exposure to the social sharing of emotion: Emotional impact, listener responses and secondary social sharing. *Eur. J. Soc. Psychol.* 27, 37–54

61 Wróbel, M. and Imbir, K.K. (2019) Broadening the perspective on emotional contagion and emotional mimicry : The correction hypothesis. *Perspect. Psychol. Sci.* DOI: 10.1177/1745691618808523

62 Aral, S. *et al.* (2009) Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks.pdf. *PNAS December* 22, 21544–21549

63 Koban, L. and Wager, T. (2016) Beyond conformity: Social influences on pain reports and physiology. *Emotion* 16, 24–32

64 Rosenbusch, H. *et al.* (2018) Multilevel emotion transfer on YouTube: Disentangling the effects of emotional contagion and homophily on video audiences. *Soc. Psychol. Personal. Sci.* DOI: 10.1177/1948550618820309

65 Shalizi, C.R. and Thomas, A.C. (2011) Homophily and contagion are generically confounded in observational social network studies. *Sociol. Methods Res.* 40, 211–239

66 Barsade, S.G. (2002) The ripple effect: Emotional contagion and its influence on group behavior. *Adm. Sci. Q.* 47, 644–675

67 Bösch, K. *et al.* (2018) , Emotional contagion through online newspapers. , in *26th*

European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018, pp. 11–28

68 Bhullar, N. (2012) Self-ratings of love and fear on emotional contagion scale depend on the enviromental exontext of rating. *Curr. Res. Soc. Psychol.* 2, Article ID 1

69 Soroka, S. *et al.* (2019) Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proc. Natl. Acad. Sci.* DOI: 10.1073/pnas.1908369116

70 Cacioppo, J.T. *et al.* (1997) Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personal. Soc. Psychol. Rev.* 1, 3–25

71 Tsai, J.L. (2007) Ideal affect: Cultural causes and behavioral consequences. *Perspect. Psychol. Sci.* 2, 242–259

72 Lin, R. and Utz, S. (2015) The emotional responses of browsing Facebook: Happiness, envy, and the role of tie strength. *Comput. Human Behav.* 52, 29–38

73 Lin, H. *et al.* (2014) Emotional disclosure on social networking sites: The role of network structure and psychological needs. *Comput. Human Behav.* 41, 342–350

74 Romero, D.M. *et al.* (2019) Social networks under stress: Specialized team roles and their communication structure. *ACM Trans. Web* 13,

75 Cao, M. *et al.* (2017) A method of emotion contagion for crowd evacuation. *Phys. A Stat. Mech. its Appl.* 483, 250–258

76 Başak, A.E. *et al.* (2018) Using real life incidents for creating realistic virtual crowds with data-driven emotion contagion. *Comput. Graph.* 72, 70–81

77 Park, G. *et al.* (2015) Automatic personality assessment through social media language. *J. Pers. Soc. Psychol.* 108, 934–952

78 Doherty, R.W. (1997) The emotional contagion scale: a measure of individual differences. *J. Nonverbal Behav.* 21, 131–154

79 Marwick, A.E. and Boyd, D. (2011) I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media Soc.* 13, 114–133

80 Derks, D. *et al.* The role of emotion in computer-mediated communication: A review. , *Computers in Human Behavior,* 24. (2008) , 766–785

81 Waterloo, S.F. *et al.* (2018) Norms of online expressions of emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp. *New Media Soc.* 20, 1813–1831

82 Kwon, K.H. and Gruzd, A. (2017) , Is aggression contagious online? A case of swearing on Donald Trump's campaign videos on YouTube. , in *Proceedings of the 50th Hawaii International Conference on System Sciences (2017),* pp. 2165–2174

83 Nguyen, T. *et al.* (2014) Affective and content analysis of online depression communities. *IEEE Trans. Affect. Comput.* 5, 217–226

84 Chmiel, A. *et al.* (2011) Collective emotions online and their influence on community life. *PLoS One* 6, e22207

85 Reinecke, L. and Trepte, S. Authenticity and well-being on social network sites: A two-wave longitudinal study on the effects of online authenticity and the positivity bias in SNS communication. DOI: 10.1016/j.chb.2013.07.030

86 Hochschild, A.R. (1979) Emotion work, feeling rules, and social structure. *Am. J. Sociol.*

85, 551–575

87 Vogel, E.A. *et al.* (2014) Social comparison, social media, and self-esteem. *Psychol. Pop. Media Cult.* 3, 206–222

88 Lee, S.Y. *et al.* (2016) What makes us click like on Facebook? Examining psychological, technological, and motivational factors on virtual endorsement. *Comput. Commun.* 73, 332–341

89 Anduiza, E. *et al.* (2009) Political participation and the internet: A field essay. *Inf. Commun. Soc.* 12, 860–878

90 Castells, M. (2012) *Networks of outrage and hope : Social movements in the internet age*, (2nd edn) Polity Press.

91 Lewis, K. *et al.* (2014) The structure of online activism. DOI: 10.15195/v1.a1

92 Mauss, I.B. *et al.* (2005) The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion* 5, 175–190

93 Mehdizadeh, S. (2010) Self-presentation 2.0: narcissism and self-esteem on Facebook. *Cyberpsychology, Behav. Soc. Netw.* 13, 357–364

94 Barash, V. *et al.* (2010) , Faceplant: Impression (Mis)management in Facebook status updates. , in *ICWSM 2010 - Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, pp. 207–210

95 Schlosser, A.E. (2020) Self-disclosure versus self-presentation on social media. *Curr. Opin. Psychol.* 31, 1–6

96 Gross, J.J. (2015) Emotion regulation: Current status and future prospects. *Psychol. Inq.*

26, 1–26

97 Vuillier, L. *et al.* (2019) Amount and diversity of digital emotional expression predicts happiness. *Harv. Bus. Rev.*

98 Dennis, A.R. *et al.* (2008) Media, tasks, and communication processes: a theory of media synchronicity. *MIS Q.* 32, 575–600

99 Barasch, A. and Berger, J. (2014) Broadcasting and narrowcasting: How audience size affects what people share. *J. Mark. Res.* 51, 286–299

100 Daft, R.L. and Lengel, R.H. (1986) Organizational information requirements, media richness and structural design. *Manage. Sci.* 32, 554–571

101 McKenna, K.Y.A. and Bargh, J.A. (1998) Coming out in the age of the internet: Identity “demarginalization” through virtual group participation. *J. Pers. Soc. Psychol.* 75, 681–694

102 McKenna, K.Y.A. and Bargh, J.A. (2000) Plan 9 from cyberspace: The implications of the internet for personality and social psychology. *Personal. Soc. Psychol. Rev.* 4, 57–75

103 Misoch, S. (2015) Stranger on the internet: Online self-disclosure and the role of visual anonymity. DOI: 10.1016/j.chb.2015.02.027

104 Ma, X. *et al.* (2016) , Anonymity, intimacy and self-disclosure in social media. , in *Conference on Human Factors in Computing Systems - Proceedings*, pp. 3857–3869

105 Ho, S.S. and Mcleod, D.M. (2008) Social-Psychological Influences on Opinion Expression in Face-to-Face and Computer-Mediated Communication. *Communic. Res.* 35,

106 Back, M.D. *et al.* (2010) Facebook profiles reflect actual personality, not self-idealization.

Psychol. Sci. 21, 372–374

107 Vazire, S. and Gosling, S.D. e-Perceptions: Personality impressions based on personal websites. , *Journal of Personality and Social Psychology*, 87. Jul-(2004) , 123–132

108 Soleymani, M. *et al.* (2017) A survey of multimodal sentiment analysis. *Image Vis. Comput.* 65, 3–14

109 You, Q. *et al.* (2015) , Robust image sentiment analysis using progressively trained and domain transferred deep networks. , in *Proceedings of the National Conference on Artificial Intelligence*, 1, pp. 381–388

110 Yang, B. and Lugger, M. (2010) Emotion recognition from speech signals using new harmony features. *Signal Processing* 90, 1415–1423

111 Ribeiro, F.N. *et al.* (2016) SentiBench: A benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Sci.* 5, 23

112 Thelwall, M. *et al.* (2012) Sentiment strength detection for the social web. *J. Am. Soc. Inf. Sci. Technol.* 63, 163–173

113 Wilson, J. and Hernández-Hall, C. (2014) VADER: A Parsimonious rule-based model for sentiment analysis of social media text. *Eighth Int. AAAI Conf. Weblogs Soc. Media* DOI: 10.1210/en.2011-1066

114 Socher, R. *et al.* (2013) , Recursive deep models for semantic compositionality over a sentiment treebank. , in *EMNLP 2013 - 2013 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, pp. 1631–1642

115 Maynard, D. and Bontcheva, K. Challenges of evaluating sentiment analysis tools on

social media. , in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016). 10th edition of the Language Resources and Evaluation Conference*, pp. 1142–1148

116 2015 Emoji Report. at <emogi.com/documents/Emoji_Report_2015.pdf>

117 Derks, D. *et al.* (2008) Emoticons in computer-mediated communication: Social motives and social context. *Cyberpsychology Behav.* 11, 99–101

118 Batson, C.D. (1975) Rational processing or rationalization? The effect of disconfirming information on a stated religious belief. *J. Pers. Soc. Psychol.* 32, 176–184