A longitudinal study of emoticon use in text messaging from smartphones

Chad C. Tossell [a,*], Philip Kortum [a], Clayton Shepard [a], Laura H. Barg-Walkow [b], Ahmad Rahmati [a], Lin Zhong [a]

[a] Department of Psychology, Rice University, 6100 Main Street, MS-25, Houston, TX 77005, United States
[b] Human Factors & Aging Laboratory, Georgia Institute of Technology, 654 Cherry St., Atlanta, GA 30332-0170, United States

ARTICLE INFO

ABSTRACT

Our goal in the present study was to understand how emoticons are used in text messaging and, in particular, how genders differed in the frequency and variety of emoticons used via this medium. Previous research has found small and sundry differences in emotive expression online suggesting that technology has closed the gender gap. However, the data collected in these studies were public. In this study, we collected real portions of private communications data from individuals' smartphones over a 6-month period. SMS messages, in general, were not used very much overall, with only 4% of all messages containing at least one emoticon. Still, differences between genders manifested in the amount and variety of emoticons used. Females sent more messages with emoticons; however, surprisingly, males used a more diverse range of emoticons.

Published by Elsevier Ltd.

1. Introduction

Text messaging (Short Messaging Service; SMS) has become an important mode of communication throughout the world and is increasing at a rapid rate, with an estimated eight trillion text messages to be sent this year alone (Global Mobile Statistics, 2011). In general, users employ this medium to coordinate activity, maintain social relationships, fill dead time, and share information with others in their social network (Ling, 2005). For teenagers, text messaging is the most popular way to communicate with their social networks, exceeding face-to-face (F2F) communications, emails, and voice calls (Pew Internet & American Life Project, 2010).

Our interest in the current study is the use of emoticons via this modality. Similar to other types of computer mediated communication (CMC), users can include emoticons within text messages to provide socioemotional context. These visual cues have been noted as the primary way to express emotion in CMC (Riva, 2002) and a way to replace non-verbal communications when not F2F (Walther & D'Addario, 2001). Most studies that have examined emoticon use have used information found on public portals via another medium (e.g., instant messaging, Derks, Bos, & von Grumbkow, 2007). Other studies have focused on emoticon use in laboratory settings in order to better understand the influence of emoticons on interpreting messages (Derks et al., 2007; Lo, 2008; Walther & D'Addario, 2001).

By contrast, this is an examination of emoticon content within text messages obtained from users' smartphones in the wild.

As such, we seek to holistically understand how people use emoticons via this ubiquitous and private method of communicating. Real communications data were collected from users' iPhones unobtrusively over the course of 6 months. These data were analyzed to characterize the frequency and variety of emoticon usage through the SMS channel. We also build on previous research and assess gender differences.

1.1. Background

Communication is not just a matter of speaking, writing, and interpreting words. Indeed, it is a complex process that involves factors such as content, language, grammar, experience, and nonverbal cues (Rezabek & Cochenour, 1998). Many researchers have noted the importance of nonverbal cues to understanding the meaning and nature of the message in F2F (Argyle, 1988). However, communication theories (e.g., social presence theory) have purported that CMC lacks contextual information and that the medium is disruptive for understanding the content and nature of messages (Sproull & Kiesler, 1986; Walther, 1992). For instance, this lack of contextual information has been blamed for causing electronic message recipients to perceive the senders of those messages as behaving rudely and offensively (Jenson, 2005).

More recent work, however, has shown that emoticons can provide this information and enhance CMC (Derks, Fischer, & Bos, 2008). Walther and D'Addario (2001) defined emoticons as graphic representations of facial expressions that are embedded in electronic messages. These often include punctuation marks and

* Corresponding author. Address: 6100 Main Street, MS-25, Department of Psychology, Rice University, Houston, TX 77005, United States. Tel.: +1 713 348 5038; fax: +1 713 348 5221.
E-mail address: chad.tossell@rice.edu (C.C. Tossell).

letters to create expressions such as happy, sad, or frustrated (which appear :), :(, and :/ respectively). Many researchers have suggested these cues enhance written communication in the same way visual or body language supports verbal communication (Derks et al., 2008; Rezabek & Cochenour, 1998). When studied empirically, viewing text online without emoticons led to incorrect interpretation of the nature of the message and the senders' attitude (Lo, 2008). The inclusion of emoticons helped readers better understand the level and direction of the emotional context surrounding the message relayed over the internet.

Rezabek and Cochenour (1998) analyzed emails on listservs for emoticon content and frequency of use. Emoticons were used in 1–25% of the emails compiled from various listservs. According to the authors, many factors that could have influenced the large variance across emails were not assessed (e.g., social tie strength, gender, age, location, etc.). The listserv with the most messages ($N = 349$) consisted of only 6% that contained at least one emoticon.

Emoticons are used more often in synchronous communications (Derks et al., 2008). In instant messaging, positive emoticons were used at a higher frequency than negative emoticons and the use of emoticons strengthened the valence of the message (Lo, 2008). Emoticons were primarily used to express emotion, strengthen messages, and display humor or sarcasm. In studies of distributed teamwork, users regularly opted to use emoticons in team communications when such emoticon utility was made available. In situations where the teams did not have emoticon use available, users were not as satisfied with the system used to complete the artificial task in the laboratory (Rivera, Cooke, & Bauhs, 1996).

Because of the brevity of SMS communications and the fact that it is used both synchronously and asynchronously (Kasesniemi & Rautianen, 2002), it is possible that emoticon usage may have different patterns of use or enhanced importance. Ling (2005) examined 882 messages using phone interviews to gather data and found that only 6% of these messages contained emoticons. In other survey-based research (Qiao, 2010), 88% of a Chinese sample used emoticons. These users preferred SMS over other media and F2F communications to express emotions to others in their social networks. Findings also showed that these users largely used emoticons in SMS for humor and as a substitute for non-verbal communication.

Since females use more non-verbal communication in F2F encounters (Derks et al., 2008), researchers have been interested in understanding if this is also true in CMC. Interestingly, studies focused on gender differences have yielded mixed results. On the web (Wolf, 2000), males did not use many emoticons on sports newsgroups where most other viewers are also male. However, when males joined mixed-gender groups, they used emoticons more frequently. The authors suggested that both males and females sought to clarify emotional states when both genders were viewing their content. In addition, though the frequency of emoticons found on these mixed-gender forums were roughly equivalent, males and females used them differently. The former used them for humor and to display emotions while the latter used them mostly for sarcasm. A similar gender-use pattern has been found on IM (Lee, 2003).

Emoticon use for different kinds of tasks (task-oriented and social-oriented) has failed to show gender differences (Derks et al., 2007). While emoticons are used more in socially oriented tasks overall, males and females both use emoticons at the same rate. In another study (Huffaker & Calvert, 2005), content on blogs was examined for gender differences. Contrary to the popular typecast, male weblogs consisted of more emoticons compared to female weblogs.

Clearly, across all mediated communication methods, there are not static gender differences (Derks et al., 2008). Emoticons are used differently across tasks, contexts, and mediums. Similar to facial expressions and other non-verbal communications, emoticons are helpful to communicate social cues, emotion, and clarify the meaning of the message. How often they are really used in SMS as well as gender differences in patterns of usage are empirical questions we address in the present research.

1.2. Hypotheses

Given the importance of emoticons in providing socioemotional context, we expected a large percentage of messages would contain at least one emoticon. The research on gender differences in emoticon use is mixed. However, many of the studies that found no differences either used contrived tasks or analyzed public content. Since text messages are mostly private communications with friends (Hakkila & Chatfield, 2005; Ling, 2005), we expected that our results would most closely resemble research found in F2F communications, with females sending more emoticons in their text messages compared to males. We also expected females to use a wider variety of emoticons compared to males.

2. Method

This study applies a quasi-experimental approach using naturalistic and longitudinal data to better understand the amount and variety of emoticons used in text messaging as well as gender differences. Real communications data were collected automatically from users' iPhones for a period of 6 months. Since text messages are considered extremely private (Hakkila & Chatfield, 2005), we obfuscated the textual data and only recorded the emoticons. For extensive details on the methodology used in this study, see our previous report (Shepard, Rahmati, Tossell, Zhong, & Kortum, 2011). We believe that the current study represents the first naturalistic and longitudinal study that collects real emoticon use from text messages in the wild.

2.1. Participants and materials

Twenty one students were recruited to participate in the research. These students had diverse academic majors, socioeconomic levels, and ethnicities. Eleven subjects were male and 10 subjects were female. iPhones that ran iOS 3.1.3 were provided to each subject free of charge over the 6 month study period. A custom logger, which operated as a background process and did not interrupt usage, was installed on each iPhone. Data were automatically captured every night and did not require user actions to record.

2.2. Procedure

Participants were required to use the instrumented iPhones as their primary mobile phone. Phones were distributed to participants at the same time. Outside of this encounter, only one other meeting was scheduled in this timeframe to collect survey information. The low number of meetings was intentional to decrease participant reactivity.

2.3. Data collection

The data collected included all text messages sent and received with the timestamp, number of words, and emoticons (number and type) within each message. Privacy was maintained by anonymizing users and the obfuscating of word content. Every subject signed an informed consent statement before they received iPhones. Emoticon type, length, and emoticon markers such as : and – were used to identify emoticons in text messages.

3. Results

A total of 158,098 text messages were sent and received by our 21 participants over the 6 month study period. In data exploration we found one outlier that consumed 20% of this overall SMS use (i.e., he sent and received over 34,000 text messages). Since this amount was well beyond 3 standard deviations of the mean, we removed him from our analysis. However, we do describe his data later. We used previous literature (Miklas et al., 2007) to define contacts (i.e., people encountered by our participants via SMS) as friends or strangers. The former were those encountered at least 10 days of the 170 days that made up the study period. The latter were contacts encountered less than 10 days over this same period. As shown in Fig. 1, most messages were sent or received from a small subset of people within individual's social networks.

Only 4.24% of all messages contained emoticons either sent or received by our participants (M = 481.09, SD = 409.23). Of these messages, 96% contained a single emoticon and 3% contained two emoticons. Approximately 1% of all messages sent and received by our participants contained three or more emoticons (max = 9).

Table 1 shows basic descriptive statistics organized by gender. Both males and females sent more messages from their iPhones than they received. Similarly, males and females sent messages to a similar number of contacts (i.e., people) considered friends according to our definition. Females sent more messages overall and to their friends compared to our male participants. Additionally, females sent out almost two times the amount of emoticons compared to males. This distinction between males and females was even greater for messages received.

Were these differences in the volume of emoticons simply due to the larger number of messages sent? To account for these differences and differences due to verbosity, emoticon-to-message and emoticon-to-word ratios were developed and used as dependent variables. Emoticon-to-message ratios represent the number of emoticons sent per message. Emoticon-to-word ratios reflect the number of emoticons sent by the number of words used. Thus, these measures normalized the data for number of messages sent and verbosity of those messages. There was no relationship between the number of messages sent and emoticon-to-message ratio (r = −.08, p = .78) or emoticon-to-word ratio (r = −.05, p = .79).

Fig. 2 shows longitudinal differences between genders in the proportion of messages sent that contained at least one emoticon. A 2 (Gender) × 6 (Month) analysis of variance (ANOVA) was performed to assess these differences. A main effect was not found

Table 1
Overview of descriptive statistics from the data collected over a period of 170 days. Means and (standard deviations) are provided.

	Male	Female
Number of text messages sent	2144.38 (*981.01*)	2896.07 (*1133.16*)
Number of text messages received	2029.73 (*902.46*)	2776.15 (*1427.39*)
Number of friends	10.41 (*6.67*)	10.40 (*7.57*)
Number of strangers	47.98 (*37.64*)	45.26 (*20.91*)
Number of messages sent to friends	2001.66, (*1004.23*)	2406.32 (*798.37*)
Number of messages received from friends	1992.88 (*831.49*)	2311.94 (*820.18*)
Words per message sent	9.20 (*3.14*)	9.95 (*2.65*)
Words per message received	9.43 (*1.45*)	9.52 (*2.95*)
Emoticons sent	241.54 (*102.31*)	422.12 (*302.60*)
Emoticons received	201.38 (*187.19*)	478.54 (*315.52*)
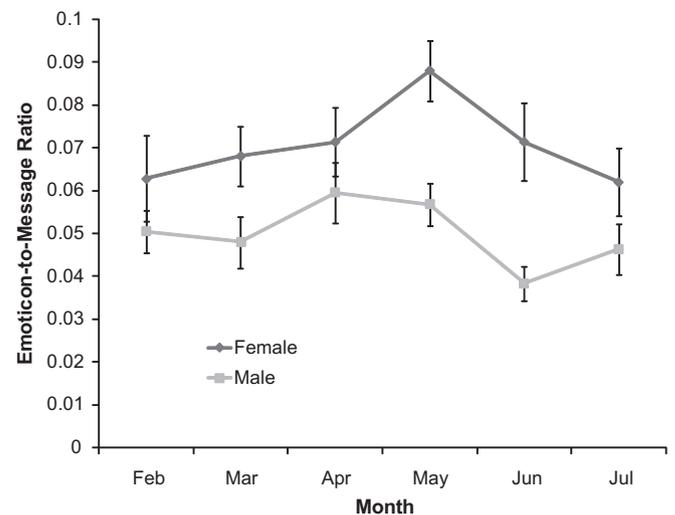
Fig. 2. Mean emoticon-to-message ratios by month and gender. Error bars represent ±1 standard error of the mean.

for month ($F(5,90)$ = 0.32, p = .84) and the interaction was also not significant ($F(5,90)$ = 1.08, p = .37). However, the main effect for gender was significant ($F(1,18)$ = 6.95, p = .02); females yielded a higher density of sent messages that contained at least one emoticon compared to males.

A similar 2 (Gender) × 6 (Month) ANOVA was conducted to assess the differences in emoticon use relative to the amount of words used over time. As shown in Fig. 3, females had a higher emoticon-to-word ratio compared to males ($F(1,19)$ = 4.86, p = .04). Neither the main effect for Month nor the interaction reached the .05 level of significance.

3.1. Variety

These gender differences were further investigated based on the type of emoticons used. A wide variety of emoticons were texted by the users in our study. Here, we focus solely on messages sent by our participants. Seventy four different emoticons were used by our participants. Only a small subset, however, made up most usage (Fig. 4). The top three emoticons (happy, sad, and very happy) made up 70% of the total amount of emoticons sent across all participants.
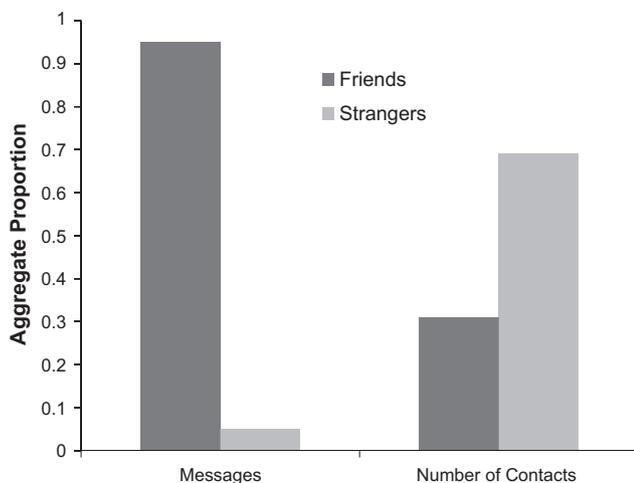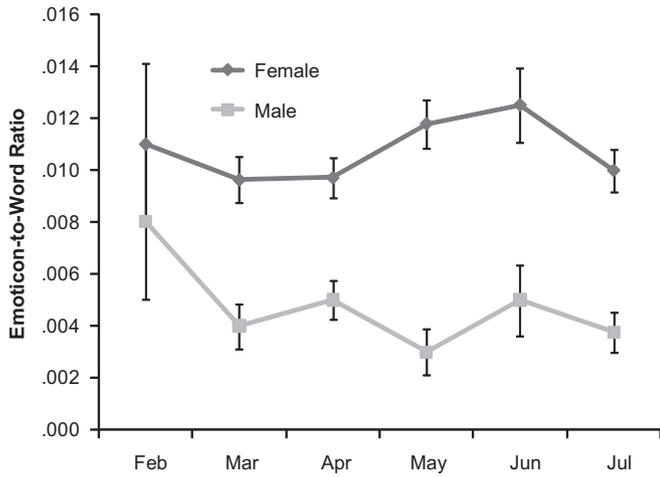
Fig. 1. Total proportions of messages sent or received by friends versus strangers and the number of people encountered that were friends and strangers.

Fig. 3. Mean emoticon-to-word ratios by month and gender. Error bars represent ±1 standard error of the mean.

Vocabulary (vocab) ratios were created for each participant. These ratios reflect the total number of unique emoticons sent divided by the total amount of emoticons sent for each participant across the length of the study. Individuals that used a more diverse range of emoticons would yield higher vocab ratios. Conversely, users that used a narrow range of emoticons repeatedly would yield lower vocab ratios. Users generally used a small subset of emoticons repeatedly ($M = 0.12$, $SD = 0.09$).

We examined gender differences on these measures across the length of the study. An independent samples t test revealed that males ($M = 0.16$, $SD = 0.10$) had higher vocab rates than females ($M = .08$, $SD = .08$) and this difference was reliable at a .05 alpha level ($t(18) = 2.12$, $p = .048$). Males, on average, sent a more diverse number of emoticons across the length of the study compared to females.

4. Discussion

This naturalistic look into emoticon use on the SMS mode of communication revealed several interesting behaviors at the aggregate level. First, in contrast to previous findings focused on other synchronous communication mediums (Derks et al., 2008), emoticons were not used very often. Over 158,098 text messages were sent and received by our 21 participants and only 4% contained emoticons. This was unanticipated given the importance of non-verbal communication in F2F communication (Ekman & Friesman, 1969; Lee, 2003) and that previous empirical work has established that the use of emoticons can help replace this information (Derks et al., 2008; Lo, 2008).

The type of communication that is done via SMS does not appear to require as much socioemotional context as other means of mediated communications, perhaps due to its simplicity and reliability (Barkhuus, 2005). Text messaging affords briefer communications and removes some of the pressures that are inherent in other types of communication (e.g., saying the right thing). Or, as (Jenson, 2005, p. 9) put it:

> When SMS first arrived, most people felt it was a trivial product. However, it met a deep need to communicate simply, less intensely, and in a time shifted manner, all of which enabled people to communicate in ways they wouldn't have normally done before.

Our study shows that this includes a lack of socioemotional cues in the form of emoticons at an even smaller percentage than obtained from previous survey reports (Ling, 2005). Most text messages were sent to a small set of people (31% of the people contacted by our users made up over 95% of the messages sent) and it appears that senders did not have the need to display their feelings in these interactions with their social networks. When emoticons were used, the frequency of use did not correlate with the number of messages sent to contacts. Thus, it does not appear that differences in emoticon use reflect the number or strength of social ties.

However, there are gender differences in the use of emoticons via this type of CMC. Previous work has suggested that technology has closed the gender gap in behaviors such as communication (Beniger, 1987). Our findings show that text messaging seems to follow similar patterns as F2F communication (Buck, Miller, & Caul, 1974). That is, females are more emotionally expressive and use more non-verbal cues compared to males on SMS which corresponds to their more frequent non-verbal displays in F2F communication (e.g., Witmer & Katzman, 1997). We found that females consistently sent messages with emoticons more than their male peers. This manifested in differences in overall amount used, emoticon-to-message density, and emoticon-to-word density. Thus, when accounting for the number of messages sent, a higher proportion of females' messages contained at least one emoticon. Females also yielded a higher concentration of emoticons relative to words compared to males.
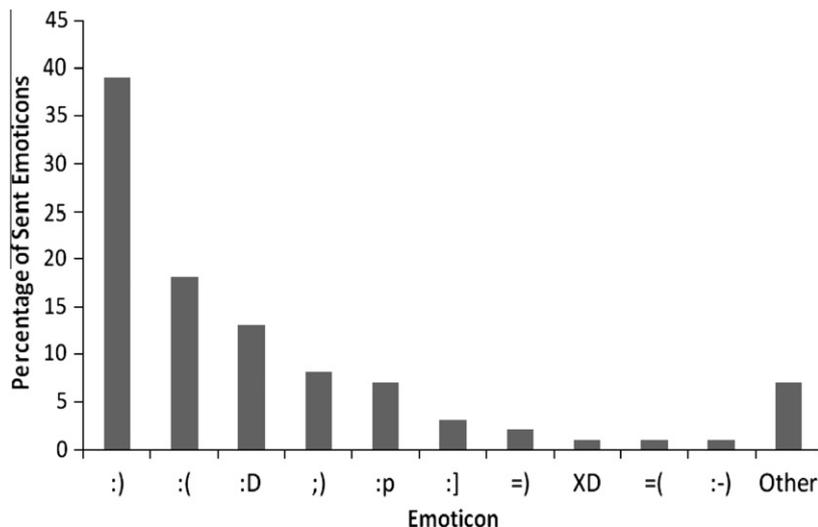

Fig. 4. Percentage of sent emoticons by the type of emoticon.

We suspected one reason for these differences could be due to the wider variety of reasons women use emoticons compared to males (Wolf, 2000). However, this did not manifest in the number of diverse emoticons used. Surprisingly, males used a wider vocabulary of emoticons in their messages. Males appeared to use emoticons for a wider range of emotions via the SMS medium. Females used a more narrow set of emoticons very frequently. Of course, the smaller set of emoticons used by females could be sent for a wider variety of purposes. Our data cannot be used to understand the intent behind the messages or the meanings of the interpretation; however, this seems like one fruitful area for future research. Additionally, factors related to the audience, setting and purpose for communicating could be explored vis-à-vis emoticon use.

Some of the most intriguing findings are the extreme behaviors manifested by the outlier mentioned above. This particular participant sent almost seven times as many text messages as the next highest SMS user (compare 34,222 to 5536 text messages). This averages out to 201 text messages sent and received for each day of the study. His rate of emoticon use was slightly lower than his male peers in terms of messages sent (.04) and number of words (.002). Clearly, there are high individual differences even amongst college students located at the same university, similar in age, race, and smartphone experience.

5. Conclusion

Many of the previous studies have focused on communications that were meant for the public eye. SMS messages are inherently and extremely private. Our naturalistic peek into these private messages showed a lack of emoticon use overall, with only 4% of all messages containing at least one. Still, differences between genders manifested over a period of 6 months even with our small sample. These gender differences were obtained from participants' real communications and provide a first look into the nature of this oft-used medium outside of self reports.

Of course, this research has only provided a glimpse into the complex nature of real mediated communications. We hope it promotes additional inquiry, away from laboratories, to understand the complexities of human communications through technology.

Acknowledgements

References

Argyle, M. (1988). *Bodily communication*. London: Methuen.
Barkhuus, L. (2005). Why everyone loves to text message: Social management with SMS. In *Proc. of group '05*. New York: ACM.
Beniger, J. (1987). Personalization of mass media and the growth of pseudo-community. *Communication Research, 14*, 352–371.
Buck, R., Miller, R. E., & Caul, W. F. (1974). Sex, personality, and physiological variables in the communication of affect via facial expression. *Journal of Personality and Social Psychology, 30*, 587–596.
Derks, D., Bos, A. E. R., & von Grumbkow, J. (2007). Emoticons and social interactions on the internet: The importance of social context. *Computers in Human Behavior, 23*, 842–849.
Derks, D., Fischer, A. H., & Bos, A. E. R. (2008). The role of emotion in computer-mediated communication: A review. *Computers in Human Behavior, 24*, 766–785.
Ekman, P., & Friesman, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage and codings. *Semiotica, 1*, 49–97.
Global Mobile Statistics (2011). <http://mobithinking.com/mobile-marketing-tools/latest-mobile-stats>.
Hakkila, J., & Chatfield, C. (2005). It's like if you opened someone else's letter – User perceived privacy and social practices with SMS communication. In *Proc. of mobile HCI conference*. Salzburg, Austria.
Huffaker, D. A., & Calvert, S. L. (2005). Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication, 10*. <http://jcmc.indiana.edu/vol10/issue12/huffaker.html>.
Jenson, S. (2005). Default thinking: Why consumer products fail. In R. Harper, A. Taylor, & L. Palen (Eds.), *The inside text: Social perspectives on SMS in the mobile age*. London: Kluwer.
Kasesniemi, E., & Rautianen, P. (2002). Mobile culture of children and teenagers in Finland. In J. E. Katz & M. Aakhus (Eds.), *Perpetual contact: Mobile communication, private talk, public performance* (pp. 170–192). Cambridge: Cambridge University Press.
Lee, C. (2003). *How does instant messaging affect interaction between genders?* Stanford, CA: The Mercury Project for Instant Messaging Studies. <http://www.stanford.edu/class/pwr3-25/group2/projects/lee.html>.
Ling, R. (2005). Mobile communication vis-à-vis teen emancipation, peer group integration and deviance. In R. Harper, A. Taylor, & L. Palen (Eds.), *The inside text: Social perspectives on SMS in the mobile age* (pp. 175–189). London: Kluwer.
Lo, S. K. (2008). The nonverbal communication functions of emoticons in computer-mediated communication. *CyberPsychology & Behavior, 11*, 595–597.
Miklas, A. G., Gollu, K. K., Chan, K. W., Saroiu, S., Gummadi, K. P., & de Lara, E. (2007). Exploiting social interactions in mobile systems. In J. Krumm et al. (Eds.), *UbiComp 2007 LNCS 4717*. Berlin: Springer-Verlag.
Pew Internet & American Life Project (2010). <http://www.pewinternet.org>.
Qiao, Q. (2010). *The adoption and use of SMS among Chinese teenagers*. Memoire: Universite de Montreal.
Rezabek, L. L., & Cochenour, J. J. (1998). Visual cues in computer-mediated communication: Supplementing text with emoticons. *Journal of Visual Literacy, 18*, 201–215.
Riva, G. (2002). The sociocognitive psychology of computer-mediated communication: The present and future of technology-based interactions. *CyberPsychology & Behavior, 5*, 581–598.
Rivera, K., Cooke, N. J., & Bauhs, J. A. (1996). The effects of emotional icons on remote communication. In *Proc. of computer human interaction*. New York: ACM.
Shepard, C., Rahmati, A., Tossell, C. C., Zhong, L., & Kortum, P. (2011). Livelab: Measuring wireless networks and smartphone users in the field. *ACM Sigmetrics Performance Evaluation Review, 38*(3), 15–20.
Sproull, L., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Sciences, 32*, 1492–1512.
Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction. *Communication Research, 19*, 52–90.
Walther, J. B., & D'Addario, K. P. (2001). The impacts of emoticons on message interpretation in computer-mediated communication. *Social Science Computer Review, 19*, 324–347.
Witmer, D. F., & Katzman, S. L. (1997). Online smilies: Does gender make a difference in the use of graphic accents? *Journal of Computer-Mediated Communication, 2*. http://jcmc.indiana.edu/vol2/issue4/witmer1.html.
Wolf, A. (2000). Emotional expression online: Gender differences in emoticon use. *CyberPsychology & Behavior, 3*, 827–833.