Chapter submitted to:

Volume 16 of book series Research on Emotions in Organizations. Volume is themed "Emotions and Service in the Digital Age". The Symposium your group held at the recent past EMONET Conference was right on topic. A book chapter(s) could be used to further showcase the work your group is doing without detracting from the journal publications you intend. Editorial Assistant Mikaela Dakic.

Editor: Charmine E. J. Härtel c.hartel@uq.edu.au

Table of Contents

Opportunities, Tools and New Insights: Evidence on Emotions in Service from Analyses of Digital Traces Data

Anat Rafaeli
Galit Yom Tov
Shelly Ashtar
Daniel Altman

Technion – Israel Institute of Technology

Chapter submitted to Research on Emotions in Organizations. Volume 16
Theme of volume "Emotions and Service in the Digital Age".
Editor: Charmine E. J. Härtel

December 2019

Digital Age Opportunities, Tools and New Insights: Evidence on Emotions in Service from Analyses of Digital Traces Data

Opening: Digital traces organic data: New and unique research opportunities

Can anyone live and work in the 21st century without some digital service interactions? We buy through Amazon or Ali Express, book flights and hotels through Expedia or Booking, and communicate with service agents through chats, texts, Facebook, Twitter and email. The issue of emotion in service is simultaneously ubiquitous – one cannot begin to talk about service without having someone intervene with their most recent story about a frustrating or annoying service situation.

At the same time, the experience of emotion in service is still very abstract. What exactly happens in a service interaction? What emotions do customers or service agents feel or express? How do emotions evolve through service conversations? And what customer emotions do service agents encounter during their work shifts? The prevalence of service in modern life has brought with it increasing research attention to emotion in service. Yet despite the growing amount of research attention these and many similar issues remain a mystery. Available research of emotions in service relies primarily on self-report data (e.g., Groth and Grandey, 2012; Wirtz and Mattila, 2003), qualitative explorations or field work based on observations (e.g., Pugh, 2001), and experimental manipulations (Van Kleef, XX; Rafaeli et al, 2012). Yet digital age 21st century technologies afford new sources of data, and new approaches to data collection and analyses, and provide fascinating opportunities for new insights.

For over a decade now, technology is a key facilitator pf customer service. For a long time, service has been conducted through telephone call centers, where "calls are recorded for quality assurance" (and perhaps for legal reasons). Such recordings offered an invaluable resource for access to the actual communication between agents and customers. Such research however, traditionally relied on the labor-intensive process of transcribing the conversations and manually coding key themes (e.g., Rafaeli, Ziklik & Doucet, 2008; REF Danielle UBC paper). But increasingly, modern-day technologies afford tools for automatic recording and retrieval of the full data comprising service interactions. Traditional service media (face-to-face, telephone) are increasingly replaced with sophisticated technology-mediated encounters, which allows customers and service employees to be in different physical locations, connecting via a technological interface (Schumann et al. 2012, Froehle and Roth 2004). From a research perspective, these technologies provide a gold mine of archives of service conversations. One such development is services delivered using written messages (chats, texting, twitter), which mean that employees and customers communicate in writing. Communication can be through corporate websites, Twitter or Facebook, or through mediators, such as http://LivePerson.com, a company that sells other firms tools for text-based service communication between customers and service agents. Such Digital Age service platforms are unique in the opportunities they afford for untangling some of the gaps in our understanding of emotions in service delivery.

Another benefit of digital age service delivery is the opportunities allowing for access to direct and accurate measures of meta-data about service conversations. Not only are the full content of the service conversations accessible, they can also be matched with when the conversation occurs, how long it lasted, what else happened before or after the conversation, and more. Making the most of such data is best done in collaborations with Computer Science, Operations Research, and Data Science colleagues. Such interdisciplinary work is still rare in the OB community, and one of the goals of this chapter is to encourage researchers to pursue such collaborations, in order to explore beyond the scope our research thus far. The tools and methodologies required for capitalizing on the opportunities of new-age data are still somewhat of a challenge and a puzzle to OB researchers. But this does not mean they should be ignored. Rather, they can be embraced and learned through collaborations. In this spirit, we describe here our recent work which relied on such a collaboration. Our work only begins to show the possibilities afforded by automated tools for studying emotion through analyses of archives of digital service conversations.

Terms, Resources and Tools for Digital Service Research about Emotion

A key feature of digital service research is that the research does not begin with collection of data based on a pre-defined research design. Rather, the idea is that research is based on analyses of data that were generated in genuine service conversations between agents and customers. Such conversations are archived by the digital platforms through which they occur and can be retrieved and analyzed using relevant automated tools. We have referred to these data as ***Digital Traces*** (Rafaeli et al, 2019), because the data trace actual, spontaneous and genuine behaviors of service agents and customers. The data are continuously documented by digital devices (e.g., corporate servers, social media platforms (e.g., Twitter, Redditt, Quora). The traces of the expressions and behaviors of customers and service agent are naturally occurring data points, that can be used to study various research questions. The design of the data is not based on the research question, rather the researcher needs to identify and construct research variables from available data. Others have labeled such data as ***Organic Data*** (Xu, Zhang & Zhou, 2019), which we find a very useful term, because it emphasizes that the data are generated naturally as behaviors occur. In the study of service, organic data depict verbal and non-verbal expressions as conversations unfold. Organic data document people's spontaneous behavior, with no intervention and potential bias due to researchers' predictions or planned research design. Importantly, organic data can be extracted without too much technical training in Computer Science (cf. API Refs).

Digital traces of various elements of human behavior are already analyzed in the emerging trend called "***Computational Social Science***" (Alvarez, 2016). Social scientists are gradually joining this trend, by doing and publishing social science research using data extracted from social media, mobile phones and other digital marvels (cf. Salganik, 2018; Jones et al, 2019; Hengchen et al, 2015; Staats et al, 2017). There is also some penetration of use of somewhat similar sources of data referred to as ***Online Panel Data*** into management research (Porter et al 2018 JOM; Walter et al, 2019). Our goal of this chapter is to inform and attract researchers of emotion in service to join this trend.

A subgroup of the Computational Social Science research trend includes Computer Science and information systems researchers who study emotion. This group uses tools that have been developed to automatically detect and identify the presence of emotion, and various features of emotion in human expression and behavior. Such tools have been in development for over ten years now (cf. Wilson et al., 2005; Pang et al, 2008), and their use is rapidly growing. The growth in the number of published papers in Computer Science, Information Systems and Robotics that study "sentiment" or "emotion" is exponential (see Rafaeli et al, 2019). Importantly, this growth is not research done by psychologists or other social scientists, but rather by computer scientists who attempt to study psychology or social science questions. For example, Paltoglou et al (2010) describe an analysis of emotions in informal textual communication in cyberspace. Thelwall et al (2012) describe detection of sentiment strength in the social web. Settanni and Marengo (2015) describe emotion expressed in Facebook posts. Gratch et al (2015) go as far as claiming to study emotional dynamics, and Waddell (2016) describes algorithms that can tell managers how their employees are feeling. The question that we as psychologists and organizational behavior researchers must ask ourselves is whether such research makes us irrelevant. We don't claim it does. To the contrary, we call here for researchers of emotion in service to join the digital age by integrating digital age data and tools into our research.

A small subset of Computational Social Science includes studies of customer service. Misopoulos et al (2016) describe using Twitter archives to identify customer service experiences with the airline industry. Balducci & Marinova (2018) refer to analyses of "Unstructured Data (UD)" using novel technologies for promoting theoretical developments in customer service. In this vein, Herzig et al. (2016) analyzed Twitter interactions to identify specific emotions (e.g., anger, frustration) in service conversations of employees and customers. Hu, Xu, Liu et al (2018) describe creation of a tone-aware chat-bot for interactions with customer requests on social media. (See also Abney, Pelletier, Ford et al, 2017; Misopoulos, Mitic, Kapoulas et al, 2014). Psychological literature is beginning to report research using such data (cf., Jones et al, 2016; Jones et al, 2019). We have called (Rafaeli & Altman, 2016; Rafaeli et al. 2019), and call here again for more consideration and integration of these new data sources and related new tools into research on emotion in customer service.

A pivotal development to the computational research on emotion, is the emergence of tools enabling automated analyses of emotion***.*** The collective term ***Sentiment Analysis*** refers to an amalgam of methods which allow automated detection of emotion in text and speech (cf. Cambria, Das, Bandyopadhyay, & Feraco, 2017). Some prevailing tools include LIWC (Pennebaker et al, 2001; Tausczik & Pennebaker, 2010), SentiStrength (Thelwall, 2013), both of which are relatively inexpensive and can be easily implemented by any determined social scientist or aspiring graduate student. A more sophisticated tool – which is continuously updated to include state of the art information systems developments is technically named the "Recursive Neural Tensor Network" (RNTN) model, and more popularly recognized as "the Stanford Tool" (Socher et al, 2013).

In the remainder of this paper we report on some of our own research of emotion in service delivery that used automated sentiment analysis. We briefly describe the context and tool we used in our research, as a prelude to sharing new insights about emotion in service delivery that these analyses avail. We specifically describe work that analyzed organic digital traces data of ***chat-based service conversations*** which we obtained through a collaboration. The research is therefore of service conversations through written chats, and was done in a collaboration with LivePerson (http://LivePerson.com) , a firm that maintains a platform for text-based interactions between customers and brands. LivePerson serves18,000 business customers, who communicate with their customers through chat. The LivePerson platform facilitates 25 million service conversations a month, accumulating archives of data that, as we describe below, offer unobtrusive insights about emotions in service conversations.

Digital Analyses of Emotion in Customer Service

Available sentiment analysis tools have been developed and validated with specific types of verbal data, mostly various types of reviews, such as movie, restaurant or hotel reviews. The tools have not been tested or validated for the study of emotion in customer service. This limitation is important because there are unique features to service conversations, which require that available tools must first be adapted to the study of emotion in customer service. Without such adaptation the accuracy of automatic coding could be hampered.

A first challenge our research therefore embraced was of testing and adapting previously defined rules and guidelines to conducting sentiment analysis of service conversations. Our collaboration began with us as researchers guiding the LivePerson Data Science team how to create a protocol for detection and quantification of emotion in service text conversations, and how to assess the validity of the emotion detection protocol. This validation process, as described in greater detail in Yom-Tov et al., (2018), began with a test of the accuracy of state-of-the-art sentiment analysis tools – The Stanford tool (Socher et al, 2013), SentiStrength (Thelwall, 2013), and LIWC (Pennebaker et al, 2001) for identification of emotions in customer chats. These tests found an unacceptably low level of precision.

This low level of validity led us to examine the unique features of text-based conversations. We found that data used for developing and validating sentiment analysis tools --movie and other reviews-- typically comprise unambiguous, straightforward opinions, and logical and well thought out text. In contrast, customer service conversations often include short sentences, do not necessarily adhere to proper spelling or grammar, and ample slang, typos, and spelling mistakes. Service conversations also contain obscenities and extensive use of punctuation, symbols, emoticons and capitalization, features that express various aspects of emotion. Some work on identification of emotion in Twitter texts attempts to consider such features (cf. Gratch et. Al, 2015)), but little work has examined emotion expressed by customers or service agents.

We therefore proceeded to adapt the available tools to the unique nature of customer service texts (see Natapov et al. 2017). In brief, the adaptation included revisions of the lexicon of the emotional value of words and terms that are used (Taboada, Brooke, Tofiloski, Voll, and Stede, 2011), addition of non-verbal icons and images such as 😊, and some added features taken from Natural Language Processing (NLP), like those used in other models of emotion analysis (cf., Buechel and Hahn 2017, Strapparava and Mihalcea 2007, Thelwall et al. 2010). Our validation found valid performance of the tool across multiple types of customers and industries.

Once our tool was developed, it could then be used it to automatically analyze large samples of customer service conversations in different types of brands, and with different types of research foci. As elaborated next these analyses are unique in four ways: (1) they rely on analyses of large samples of actual expressions of customers, suggesting high external validity; (2) they are done automatically, with no human intervention, so offer high reliability, and minimal biases due to human error; (3) they provide access to new variables and analyses that previous research on emotion in customer service could not access without major investment of time and effort; (4) they provide data on different firms, time periods, service agents and customers, allowing for comparisons and insights at a level of granularity that exceeds anything done in previous research.

Studying Emotion in Customer Service by Analyzing Chat Based Service Conversations

1. *The magnitude of the data*

A large magnitude of data is the first benefit of the new data and analyses that we promote. For example, in the data and analyses we will describe below are based on data retrieved from an archive comprising 1.14 million service conversations (or some 14 million text messages). In any given study we cannot necessarily obtain or analyze such a huge amount of data, so like any other project we work with samples. But these samples are still significantly larger than sample reported in prevailing research on emotion. Moreover, the data represents multiple firms, different time frames (days, weeks, months) and multiple service agents and customers. And additional samples of data can be retrieved to replicate of test the robustness of any given finding.

The magnitude of these data stand in stark contrast to typical data sets in studies of emotion in customer service, where samples are often a few hundred at best. With digital traces data, samples can be obtained and compared with relative ease, without a stark effort of additional data collection. Moreover, the data are organic, and do not rely on the self-report of service agents or customers. Data also span wide ranges of time, and at resolutions of minutes, days or months. Thus, the data allow a focus on evolutions over time. To illustrate, we report below on patterns of customer expression of emotion over the course of conversations, during the shift of a specific service agent, as a function of time of day or day of the week. In addition, in contrast to most published research which reports on one firm or industry, the analyses we report refer to different firms and industries and include some comparisons between firms and industries. These comparisons have not been a major research goal for us but could be a focus of future research.

The variety of lenses of possible analyses allows us to unravel some fundamental issues regarding emotions in customer service, including issues that previous research constraints prohibited. For example, we report data that offer insights into what emotions customers *actually display* to service agents, as opposed to the customer emotions that agents remember or recall, which is what self-report data represent. We also report data on emotions displayed at different times of day, or different days of the week, and data on emotions expressed as service conversations unfold, from the perspective of the customer and the perspective of the service agent. Obviously, what these data show are dynamics of emotions in digital service conversations, rather than phone or face to face conversations. But these data nonetheless provide informative insights into the genuine dynamics of service conversations. Moreover, new tools are being developed as we write this chapter that can avail similar analyses of other forms of service delivery.

2. The nature of variables available in digital traces data.

The organic nature of digital traces means that they come as they are, based on decisions of the designers or the entities that maintain the digital platform from which they are extracted. A major disadvantage is that researchers have limited control over the full set of variables that they can assess. For example, demographic variables --often expected and presumed essential in reports on behavioral science research-- are not available to us in the analyses reported below. A key reason this information is not available is privacy. As Barbarro and Zeller (2006) showed, major effort and disguise of private information is essential to ensure that data of a specific person cannot be identified.

At the same time, the advantage of digital traces data is the diversity of other information that they do include. Some information is automatically available, notably information unavailable to researchers until now without major and costly coding efforts. For example, data about how long service conversations lasted, how many words were said by each party, how quickly participants (customers and service agent) responded to each other provides fine-tuned access to the nature of the service situation. In our projects, for example, we use the number of words customers express as indices of the amount of cognitive load or complexity of the customer request that an agent must handle. Similarly, we use the number of words that agents post as an index of the effort a customer situation demanded (REF Altman et al., MSOM). We also use the number of lines written in the course of a conversation as a measure of the how long a conversation took, and the number of lines written by each participant as the relative contribution of each participant.

Empirical Insights: Analyses of Digital Traces of over a million service chat conversations

Available Data

The insights that we obtained from the archives of service conversations maintained by LivePerson retain the privacy of customers, and therefore do not include any identifying information. The data included an ID code identifying (and keeping anonymous) the employee and the customer in each conversation, the number of customer words, the number of customer lines in the interaction, and an emotion score. The emotion score for each sentence in the conversation was generated by automatic analyses using the tool we had developed together with the firm as described above (see also Yom tov et al, 2016). The emotion score varies from -7 to +7. A score of zero (0) indicates *No Emotion*, presence and intensity of positive and negative emotions in each of the customer messages is coded as 1 to 7 for *Positive Emotions*, and -7 to -1 for *Negative Emotions*.

As reported below, for some of our analyses we aggregated these customer sentence-level scores to analyze emotions at the conversation-level of analysis. We define as *Positive* conversations that include at least one positive message and no negative messages, conversations as *Negative* when they include at least one negative message and no positive messages, conversations as *Multiple Emotion*, when they include at least one positive and one negative message. *No emotion* conversations are those where all the messages are assigned a score of 0 (zero).

Some of the also data included three customer satisfaction variables, which were solicited after the conversations had ended[1]: (1) ***Customer Assessment of Employee Performance****,* namely customer responses to the question "Thinking about the employee that you just chatted with, how would you rate him/her?" (1-5 scale, 1=poor to 5=excellent). (2) ***Net Promoter Score*** (NPS; Reichheld 2003), and index of customer satisfaction based on customer responses to the question: "How likely is it that you would recommend our company to a friend or a colleague?" (0-10 scale, 0=Not likely to recommend, 10 = highly likely to recommend). (3) Customer assessment of ***First Contact Resolution*** (FCR; Hart et al. 2006), assessed by means of customer responses to the question "Was your query resolved in this interaction?" (Closed response scale: Yes/No).

Figure 1 describes the scope of the data in terms of the number of words in customer messages, and the number of customer messages in service conversations.[2]

[1] These data are available only for part of the data because not all customers respond to the post-service surveys.
[2] As noted, Figure 1 describes one company – an air travel provider– and presented here as an overarching picture of the diversity in length of customer chat service conversations. This diversity is rarely mentioned or recognized in research on emotion in service, yet likely relevant to the full understanding of the complexity of service work.

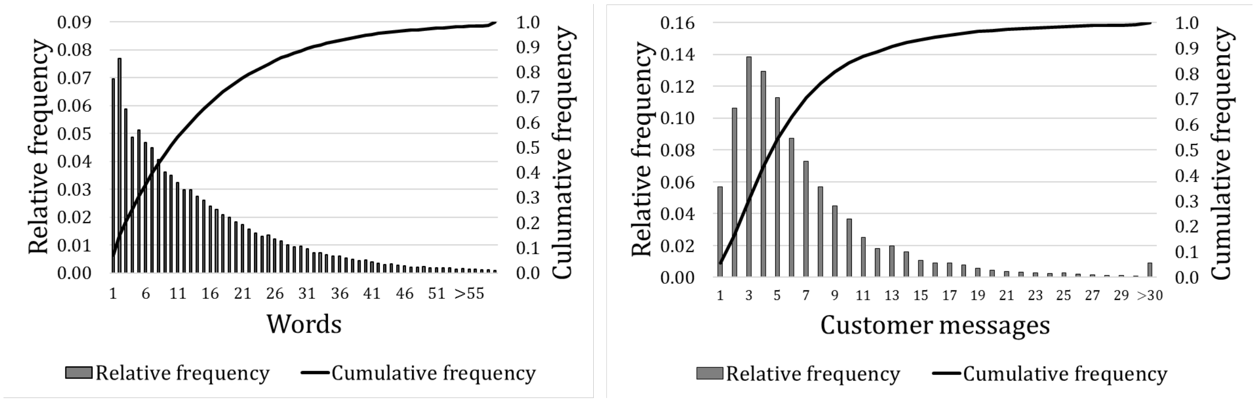


Figure 1. Number of words in customer messages and number of customer messages in service interactions, in a sample of customer interactions [Air travel firm, n= 25,714][3]

The companies in the sample were selected to represent three different industries and include a telecommunication company (n=677,936), a retail company (n=439,585), and an air travel provider (n=25,714). These samples are not huge, but they are substantially larger than typical samples in research on emotion in customer service. The data are large enough to provide insights that can are highly likely to be representative of the population, and any findings with such data likely have external validity. Moreover, the data include a lot more granularity than most other research on emotion in customer service.

To retain the privacy and anonymity, we do not know anything about the firms, the agents or the customers, and cannot report demographic information. The data does include, however, an ID code identifying (and keeping anonymous) the agent and customer in each conversation. This means we can track multiple conversations of the same customer or agent. As described below, we can therefore trace the pattern of emotion customers express over the course of conversations, or the full load of customer emotions an agent experiences over the course of a shift.

Additional statistics that these data afford go beyond number of words. Like number of words, the statistics may or may not directly regard emotion, but we believe they are relevant and should be considered in the context of research on emotion in service delivery. First, agents in the sample can converse with more than one customer at a time, up to a maximum of 3 customers[4]. During a work-shift an agent can converse with between 22 and 124 customers (Mean=64, SD=27). Second, the duration of conversations varies widely, from less than a minute to 39 minutes (Mean=11.5 minutes, SD=9 minutes). These types of data provide important insight into the interpersonally demanding nature of customer service work, and these data are not based or biased by agents' memory.

As noted, automated sentiment analyses of the text of each sentence in a chat conversation provides an emotion score for each sentence. Positive scores indicate the intensity of *Positive Emotion* (1 to 7), while negative scores indicate the intensity of *Negative Emotion* (-7 to -1). Sentences where no

[3] For brevity we do not repeat any of our reported results here for the different firms in our sample. Thus Figure 1 reports only about the data of air-travel customers. Other figures report randomly about different firms that we analyzed.

[4] Current trends develop customer service delivered through texting (SMS) rather than chat. We have begun analyzing such data, and find that there agents can converse with up to 20 customers at a given point in time.

emotion is identified in the text are defined as *No Emotion* and assigned a score of 0[5]. As reported below, these emotion scores can then be aggregated to represent the emotion in a full conversation, or to study changes in emotion over time within a conversation, or over time within an agent's work shift. Importantly, the emotion scores represent the emotions that agents and customers express. We do not have access to the emotions people feel. From a theoretical perspective, however, we can view this as direct access to the emotion that a partner to the conversation experienced; partners have access only to what the other person expressed, not to what he or she felt (Hareli & Rafaeli, 2006). Expressed emotions are therefore most likely what influences the way a conversation unfolds.

New insights from analyses of digital traces of emotions in customer service conversations

The unique data we described above can attempt to address a large range of questions. Our goal here is to illustrate the insights that such data can provide about issues and aspects of emotion in customer service that have not been previously explored. We do not report on hypothesis testing data, but rather descriptive data. As such we believe these data are like a qualitative study: They provide insights about the phenomenon being studied – emotion in service delivery – in the hope they will trigger and inspire research future research. We specifically report next (a) what emotions do customers *really* express in service conversations? (b) What customer emotions do service agents *really* encounter as they perform service delivery work? (c) How do customer emotions unfold during service conversations? (d) What are some differences between emotions expressed by customers in different industries?

What emotions do customers *really* express?

A first, somewhat surprising finding in our chat-service data regards the amount of emotion in customer behavior. Our first, simplistic look, was at the emotion expressed in individual customer messages; we found that most customer messages do not contain any emotions (see Figure 2). This finding is notable because of the common perception of emotion being so central to service delivery. Further surprising is that instances customers who do express emotion, show far more positive emotion than negative emotion! As evident in Figure 2, only 5% of customer messages express negative emotion, while 20% of messages express positive emotion. It seems that contrary to common belief the more dominant customer behavior is displaying positive not negative emotion. This is an important point because we have substantial research about the extent and impact of customer expressions of negative emotions toward service agents, and little research about expressions of positive emotions (cf. Grandey, Rafaeli, Ravid & Wirtz, 2010; Groth & Grandey, 2012; Rafaeli et al, 2012; Foulk et al, 2016).

[5] Our analyses show that a negligible (less than .10% of messages) include more than one emotion.

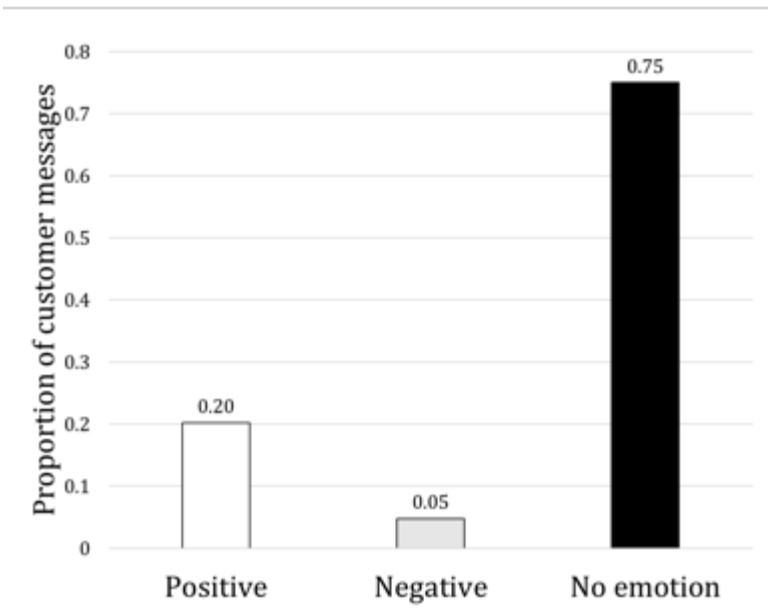


Figure 2 Emotion in Messages [*n* =3,659,053]

Figure 2. Frequencies of expressed positive and negative emotion in customer messages [Retail]6

Importantly, Figure 2 depicts a view that without analyses of digital traces data is unbearably difficult to obtain, because discerning the emotions expressed in individual lines or sentences that customers emit is an extremely labor-intensive endeavor. Reliance on self-report or on experimental manipulations provides a focused look overall emotion in service conversations, or a focus on expressions of specific emotions.

A required expansion of the analysis in Figure 2 of emotion in customer sentences is a look at the emotion throughout full customer conversations rather than individual messages. Individual messages may not convey emotions, while full conversations may exude an emotional overtone. The benefit of the analysis of digital-traces data is the relative simplicity of aggregating from individual expressions to full conversations. Thus, we continued beyond customer sentence-level emotion scores to capture the emotion overtone of full conversations. We defined conversations as having *Positive Emotion* if they include at least one positive message and no negative messages, as having *Negative Emotion* if they include at least one negative message and no positive messages, and as *Multiple Emotion*, if they include at least one positive and one negative message. When emotion was not identified in any of the messages in a conversation a conversation is defined as having *No Emotion*.

The view into emotions in full conversations -- as shown in Figure 3 -- shows that more than 80% of customer conversations do contain some expression of emotion. However, here as well negative emotions are <u>not</u> as frequent or dominant as available research seems to imply. Rather, Figure 3 shows that conversations are <u>far less likely</u> to have negative emotions, than positive emotions, multiple

[6] Since our goal in this chapter is to illustrate insights that can be gleaned using the new tools and data we describe here, we report in different figures results for different samples from different industries. Our goal is <u>not</u> to describe any specific sample, organization, or industry. Rather, our goal is to show the diverse ways that data can be extracted and described in order to trigger future research ideas. Only where our goal is to draw comparisons between samples, we provide the comparative data between the samples.

emotions, or no emotions. Only a small proportion of customers (5%) express only negative emotions, probably reflecting the 5% of customer messages that include negative emotions. In contrast, positive emotions appear in 20% of customer lines, and a dominant 57% of customer conversations have customers expressing *only positive emotions*.

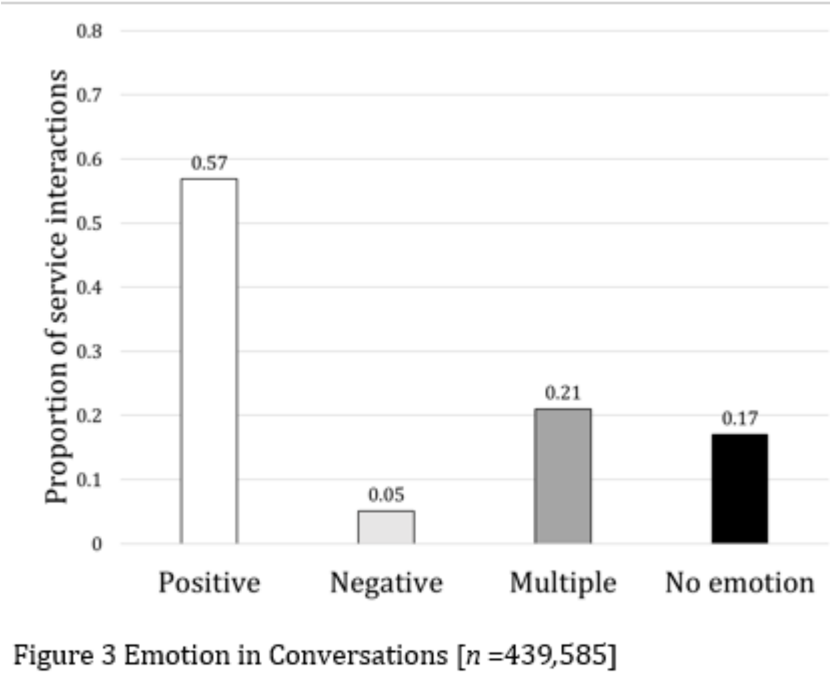


Figure 3. Frequencies of emotion in service conversations [Retail]

Another new angle that these automated analyses of digital traces of organic data provide regards the multiplicity of customer emotions. Available research frequently speaks about the presence and effects of negative customer emotion. Service agents implicitly imply that expressions of negative emotions are the frequent and critical thing in their work. The idea that negative emotions may occur on their own, or in the presence of other – positive—emotions has not been addressed in previous research. What Figure 3 shows is that a substantial portion (21%) of customers express negative emotions and other, positive emotions, within the same conversation. In other words, customers express multiple emotions, some of which are negative, others are positive. Only 17% of customer service conversations do not include any expressions of emotion, and only 5% of conversations include only negative emotions. In contrast, 57% of conversations include *only positive emotions*. This sheds a new light on customers, as more pleasant and amenable to service agents than our literature on emotion in service delivery previously presumed. And opens a host of new questions for research.

Recognizing that customers can convey multiple emotions within the same conversation raises a question about the emotional tone of a full conversation. Our analysis in Figure 3 created a taxonomy of different types of conversations. But since we can also develop a continuous score indicating the amount of positive and negative emotion expressed in each conversation. And we can continue to add up these scores, thus computing the sum of negative and positive emotion scores in a conversation, which provides a summary score for the emotional balance of each conversation. Such summing up is a crude estimate of the balance of positive and negative emotion in conversations, because it implicitly assumes that positive emotions cancel out damaging effects of expressed negative emotions by the same customer. This of course is an empirical question, that deserves research attention, but not a totally unlikely proposition. Customer expressions of positive emotions can be viewed as forms of social support or resources that customers imbued to service agents, while expressions of negative emotions are viewed as work demands and causes of depletion.

Notwithstanding, the bottom line is that such a summary score affords a way to examine the profile of service conversations in terms of the emotions that customers express. Conversations with a summary positive score suggest a conversation with an overall balance of positive emotions, while those with a summary negative score indicate a conversation that has a balance of negative emotions. The results of this computation – as depicted in Figure 4 – continue to show **more overly positive emotion conversations than overly negative emotion conversations**. The tail of positive emotion conversations in Figure 4 is much longer than the tail of negative emotion conversations. We propose this suggests more variation in the types of positive than negative emotion in a conversation. The bottom line of Figure 4 is that more than 70% of conversations where a customer expresses some emotion, have an overall positive emotion.

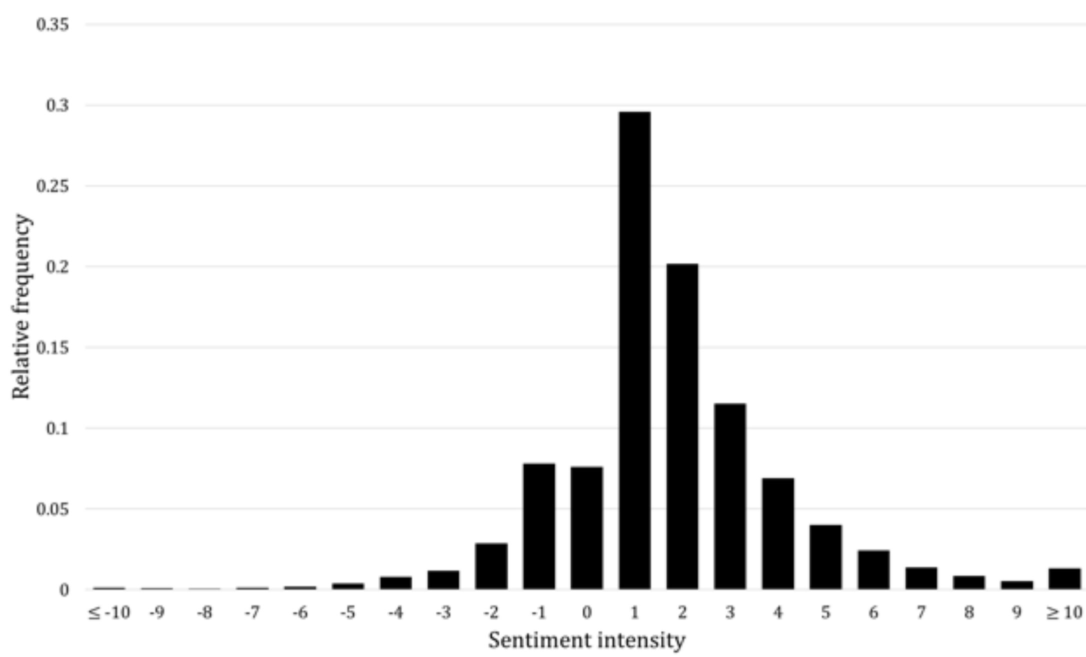


Figure 4. Distribution of the sum of intensity of positive and negative customer emotions [Transportation, *n*= 21,122]

Figure 4. Distribution of sum of positive and negative customer emotions in full conversations [Air travel, n= 21,122]

What customer emotions do service agents *really* encounter in their service delivery work?

By looking at the emotions that customers express in conversations, we obtain a more refined understanding of customers' emotions in service. For example, we can ask whether customer emotions vary by the time of day or day of the week a service conversation occurs. A complementary view, that adds more detailed insight is at the emotion customers express *within a conversation.* For example, we can look at the emotions expressed early on (at the opening of) conversations or later, toward the end of the conversations. As a start, in Figure 5 we show an analysis that focuses on emotions that customers express at the beginning and the end of their conversations. Figure 5 shows that emotions customers express do not really vary by time of day but do vary (and improve) from the first to the last message. In aggregate, customers appear to start off conversations with very mild negative emotions (sentiment scores < 0), and end conversations with expressions of mildly positive emotions (sentiment scores around 0.7). This pattern is not related to time of the day a conversation occurs.[7]

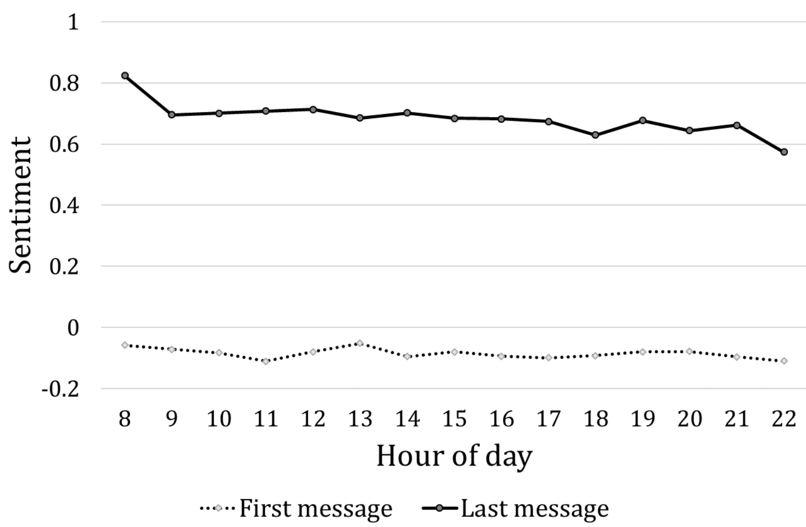


Figure 5. Emotions customers expressed in first message and last message at all hours of the day [Air travel, *n*= 25,668]

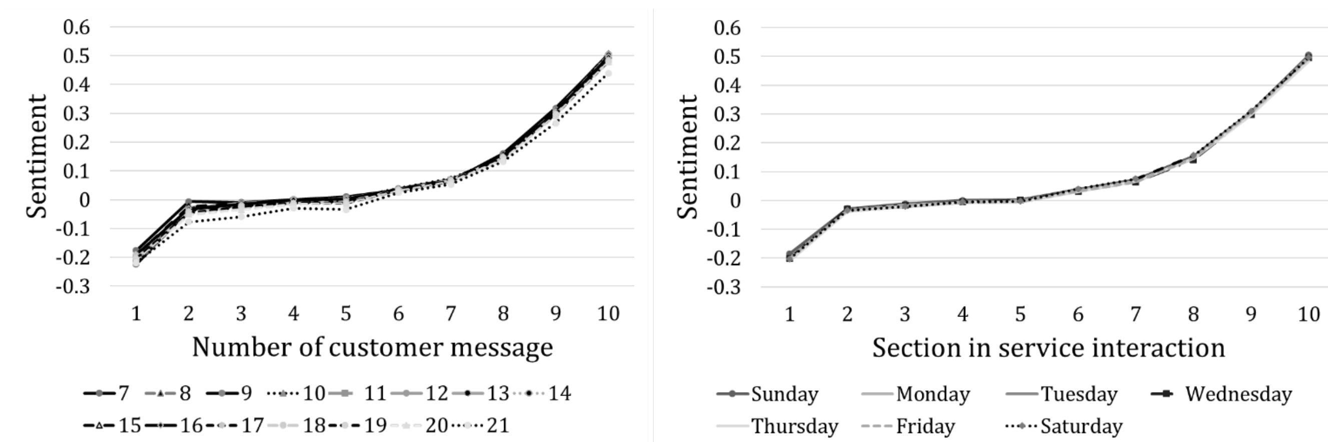
Figure 5. Customer emotion in first line and last line of a conversation with a service agent [Air travel, n= 25,668]

A more refined look at the emotion customers express within conversations is depicted in Figure 6, which shows the emotions typical to multiple stages or sections within conversations. This type of analysis portrays service encounters as a sequence of emotional displays, extending beyond the focus on peak emotions, as suggested by Verhoef, Antonides, and de Hoog, (2004), for example. Since customer conversations vary in the number lines they comprise, we must first create a standardized metric that allows comparisons of different conversations. We obtain such standardization by splitting all conversations into 10 roughly equal sections; this standardization means that sections in different interactions may comprise a different number of messages, but all conversations comprise exactly 10 sections (or 10 quartiles). With such standardization we could then average the emotion scores of all customer messages in each section and obtain a metric of the valence of customer emotion

[7] Figure 7 seems to suggest a minor effect of a drop in positive customer emotion in the end of conversations emotion in the first and last hour of the day: slightly higher positive emotion is expressed at 08:00 than at 09:00, and slightly lower positive emotion is expressed at 21:00 than at 22:00. These are not big differences and our analyses could not identify why they occur. Further and future research might be able to unravel these dynamics.

in this section. Each conversation is thus defined as comprising 10 sections, and 10 emotion scores. The result of this standardization allows us to depict the flow of emotion over the course of multiple interactions, as shown in Figure 6[8].

Figure 6 suggests that conversations can be proposed to have a standard structure comprising three stages within conversations —opening, middle (or main), and closing. Figure 6 shows a replication of the analysis of emotion by section at different hours, and different days of the week, and shows negligible variations. Similar to Figure 5, Figure 6 again suggests that conversations open with minor negative emotions, and end with more positive emotions. The main and middle of conversations shows customers as mostly neutral (non-emotional).



(a) By hour of day

(b) By day of week

Figure 6. Customer emotion expressed in different sections of service conversations, at different hours of the day and different days of the week [Telecommunication, n= 390,438]

It is not surprising that customers express negative emotion at the starts of their conversations, since initiating a service conversation means the customer has a problem. The more positive emotion expressed toward the end of conversations presumably suggests that the conversation helped resolve the customer's problem. The middle sections, where there seems to be little expression of emotions by customers, the conversations likely focus on the technical issues that the customer problem or its solution comprise. In further analyses—reported below- we examine the trajectory of the patterns of improvement of emotion, as they relate to the quality of resolution of the customer problem. We show this by connecting the pattern of improvement in emotion to post-conversation surveys of customer satisfaction.

What customer emotions do service agents encounter over the course of a work shift?

Agents often report encountering a lot of hostile customer emotions, but to our knowledge the extent to which this is an accurate depiction of service work hasn't been objectively verified. Our data allow addressing this question in two ways. First, as depicted in Figure 7, we tracked the customer

[8] The data in Figure 6 describe only the subset of conversations that included 10 or more customer messages. We repeated this analysis on all the data, including shorter interactions, as a robustness check, and found a similar pattern. In this repeated analysis, we stretched short interactions that include less than 10 customer messages, by duplicating missing quantiles. For example, for an interaction with length 5: 1,2,3,4,5, the 10 points were 1,1,2,2,3,3,4,4,5,5.

emotion encountered by a random employee on a random workday. Thus, Figure 7 is a plot of the positive and negative emotions included in the 124 conversations one random agent had during a random day.

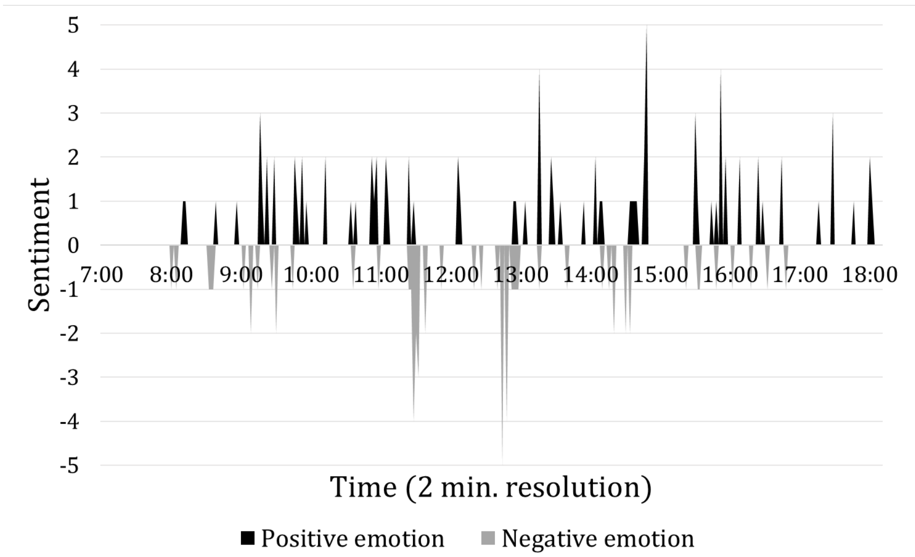


Figure 7. Customer emotion encountered by one employee during a workday [within agent analysis, n= 124 conversations]

Looking at the emotions encountered by one agent – as shown in Figure 8 -- can run the risk of a sampling bias. Perhaps we randomly selected a particularly problematic agent? Thus, we also compute an aggregation of emotions that all customers convey over the course of a full workday across all agents. This depiction removes the concern of a sampling error, an outlier or special case agent. Figure 8 thus shows the cumulative positive and negative emotions expressed by all customers over the course of a full workday. The picture depicted by Figure 8 shows a rise and fall of expressions of positive and negative emotions. The general picture is of more expressions of positive emotions than negative emotions, with some points where this is not the case. More broadly, both Figure 8 and Figure 9 suggest that service agents must handle a continuous transition between customers expressing positive emotions and customers expressing negative emotions. We propose that these transitions are the most difficult part of service agents' work, similar to the depleting and debilitating social influence that Rafaeli & Sutton (1991) described of encounters with emotionally contrasting social expressions by conversation partners.

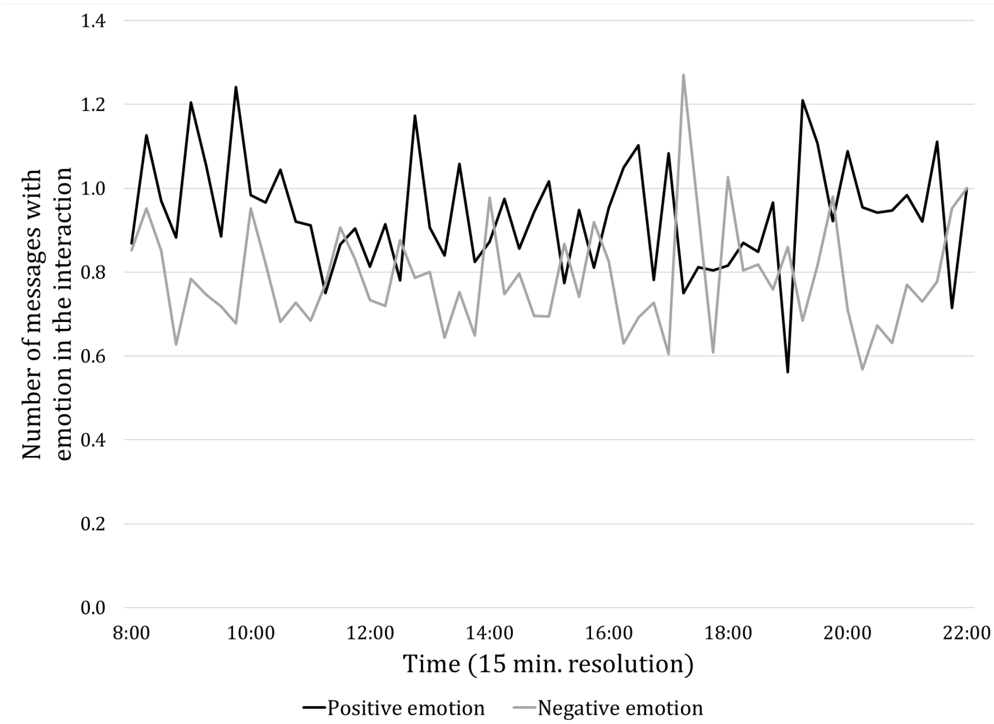


Figure 8. Total (sum of all) positive emotions and negative emotion that all customers express over the course of a workday [Air travel, n= 11,591]

Relating customer expression of emotions to customer evaluations of the service conversation.

Our analyses thus far have been purely descriptive, showing the emotions that customers express in the context of customer service conversations. An additional perspective that research relying on digital traces data can provide is the relationship of customer emotions expressed *during service conversations* to customer evaluations of the service agent and service conversation *after the service ended.* For this perspective we integrate the analyses of customer expressed emotion with data we have regarding customer evaluations of service quality. The firm we work with, like many service providers, follows up on service conversations with a text message asking customers to respond to a survey assessing their satisfaction with the service they received. Responses to such post-service surveys, however, depend on the customer's will, hence we do not have these data for all the customers in our samples. Notwithstanding, the sample sizes of our analyses here are still substantially larger than most sample sizes in previous emotion-in-customer-service research.

First, Figure 10 depicts the pattern of emotion expressed customers by *within their service conversations*, broken down by the customers' response regarding their level of satisfaction with the performance of the service agent *during the conversation*. As noted, the evaluations of the agent performance were given after the service conversation had ended. In contrast, the assessment of the expressed emotion was conducted *during the conversation*

The frame of reference in Figure 9—the bold line—depicts the mean emotion throughout the conversation in the population of customers who did not respond to a post-service survey. What the Figure 10 shows is that the emotions expressed by customers who were extremely satisfied with the agent performance climbed higher, to include expressions of more positive emotions. In contrast, emotions expressed by customers who were dissatisfied with their agent's performance (rating of 1) remained low throughout the conversations. While these customers may be argued to have started with more negative emotions from the outset of the conversation, we can see that customers who started out at the same level of negative emotions as most other customers and ended up reporting discontent with the agent performance also remained with more negative emotions than the general population of customers throughout the conversation.

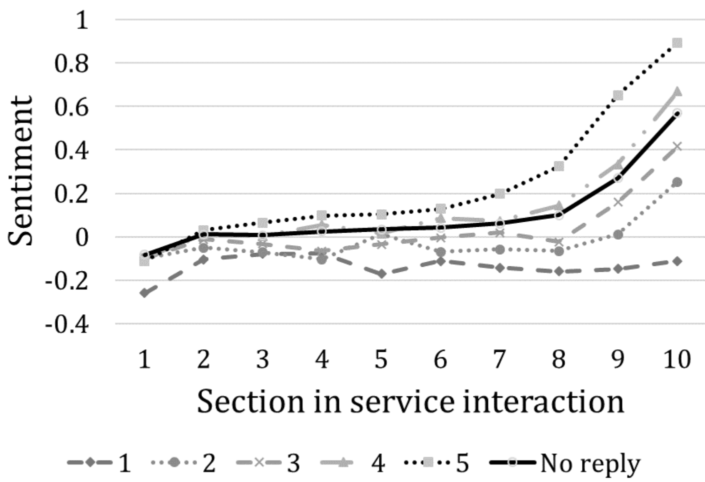


Figure 9. Customers' expressed emotions during service conversation by level of customer satisfaction with the performance of the service agent after the service conversation had ended [Air travel, n= 6,973]

A different, but related measure of customer satisfaction-- known as Net Promoter Score (NPS)—tells a similar story, as depicted in Figure 10. Figure 10 shows that, relative to the general population of non-responding customers, customers who are highly satisfied, and eager to recommend the service provider to friends or acquaintances, show a steeper improvement in the positive emotions they express. Moreover, customers who end up highly satisfied with the service they received (NPS=5) start expressing higher positive emotions earlier in the conversation than customers who were mildly satisfied (NPS=3), and customers who were not at all satisfied 9NPS=1) do not express any positive emotions during the conversation.

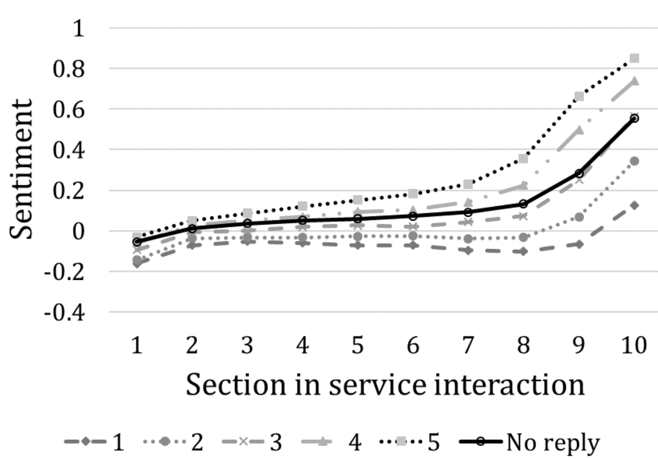
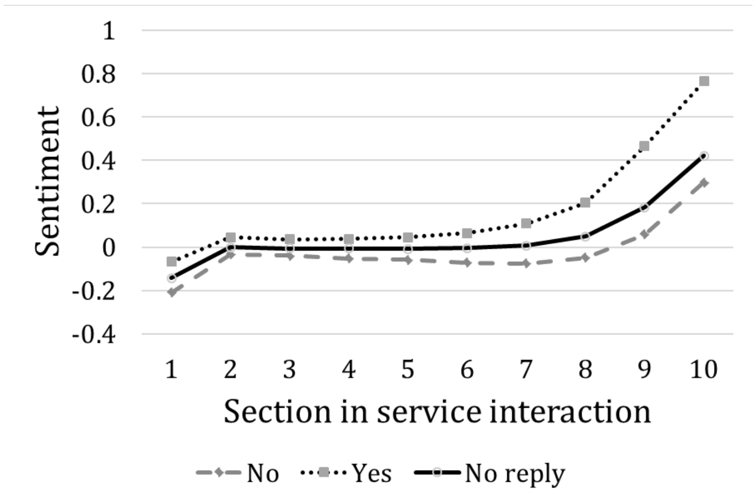


Figure 10. Customers expressed emotions during service conversation by level of customer overall satisfaction with the service agent as measured by the Net Promoter Score (NPS) [retail, n= 123,554]

Finally, in Figure 11, we show the evolving customer emotions during conversations as they relate to customer evaluations of the extent to which their service issue or problem were resolved by the conversation. This criterion – referred to in the service industry as First Call Resolution (FCR) – is an index of efficiency of handling customer service issues. The picture depicted in Figure 12 suggests that service conversations that are more effective at resolving customer needs also evoke positive customer emotions earlier in the service conversation.



(c) Customer First Contact Resolution (FCR) [Telecommunication, *n*= 390,438]

Figure 11. Customer emotion in sections of interaction by customer evaluations of whether their service issue had been resolved.

Are there differences in customer emotions between different industries?

The methods we describe – of automated analyses of emotions of digital traces of expressions of emotion by customers – are unique in the relatively easy access they provide to quick and relatively easy—automated rather than manual--analyses of large amounts of data. A new and interesting opportunity such data afford is examination and comparisons of emotions expressed by customer of different companies or industries. The data we describe here included service conversations of three firm, in three different industries: A firm providing telecommunication services, a firm providing retail services and a firm providing air travel services. We intentionally selected these three firms because they provide distinct types of services (cf., Wirtz and Lovelock 2016).

Figure 12 shows the comparison of the customer emotions in each of these firms. Figure 13 essentially replicates the distribution shown in Figure 2 for the three industries.

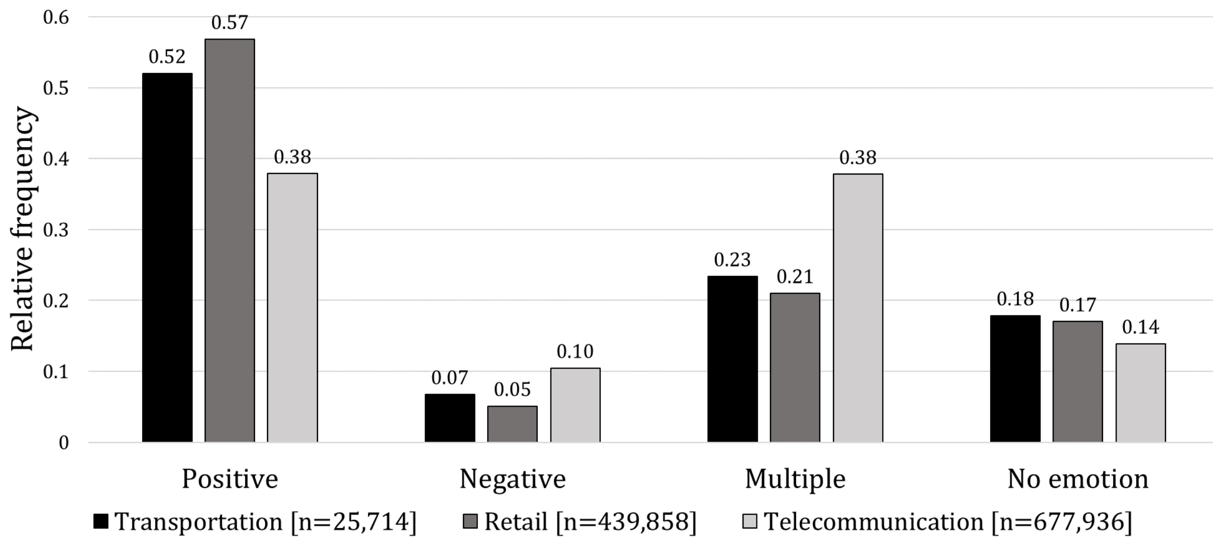


Figure 12. Customer emotion in service conversations in three firms from three industries

Figure 12. Customer emotion in service conversations in firms representing three industries

The pattern in Figure 12 shows some similarity between the three firms: less than 20% of conversations in all firm contain no emotion, and positive emotion has the largest presence in all firms (38% or more vs. 10% or less negative emotion). A Chi Square test of independence confirms a significant difference between the (substantially fewer conversations with only positive emotions in) the telecommunication firm than in the other two firms (38% compared to 52% in air travel and 57% in retail). A significant differences difference was also confirmed with more conversations with negative emotions in telecommunications than in the other two firms (10%, vs. 7% in air travel and 5% in retail). Such comparisons add to our understanding of the difficulties of handling customer support services in different firms and industries.

Beyond emotion, the data about the three firms also allows comparisons of other aspects of service delivery. We find, for example, that the average number of customer messages in a full conversation is longer in the telecommunication firm (14.8 (SD=13.2)), than in the retail firm (8.32 (SD=7.80)), and the air travel firm (8.12 (SD=7.78)). The conversations in the telecommunication firm are, on average, almost double in length than the conversations in the retail and air travel firm. Moreover, employees of the **telecommunication** industry interact with 13 to 69 (Mean=35, SD=17) customers during a shift, and their conversations vary in duration between 3 minutes and 49 minutes (Mean=19, SD=15). Customer service agents in the **retail** industry firm interact with 11 to 82 (Mean=46, SD=20) customers during a shift, and their conversations vary in duration between 3 minutes and 37 minutes (Mean=14; SD=12).

Relating customer expression of emotion to service agents' behaviors.

A particularly intriguing vantage point that our analyses can provide regards the relationship between the emotions expressed by customers and the behaviors of service agents. This question was the foundation of the early work on emotional labor, which presumed that the emotions displayed by service agents is critical to the satisfaction of customers. More recent analyses, however, posited reverse causality, showing that expressions of anger by customers can hamper employees' cognitive abilities (Rafaeli et al., 2012). In this vein, we sought to examine the relationship between customer expression of emotion and work behaviors of service agents. This line of our work is reported in greater extent in Altman et al. (2019). A key and intriguing graphic depiction of our findings is shown in Figure 13.

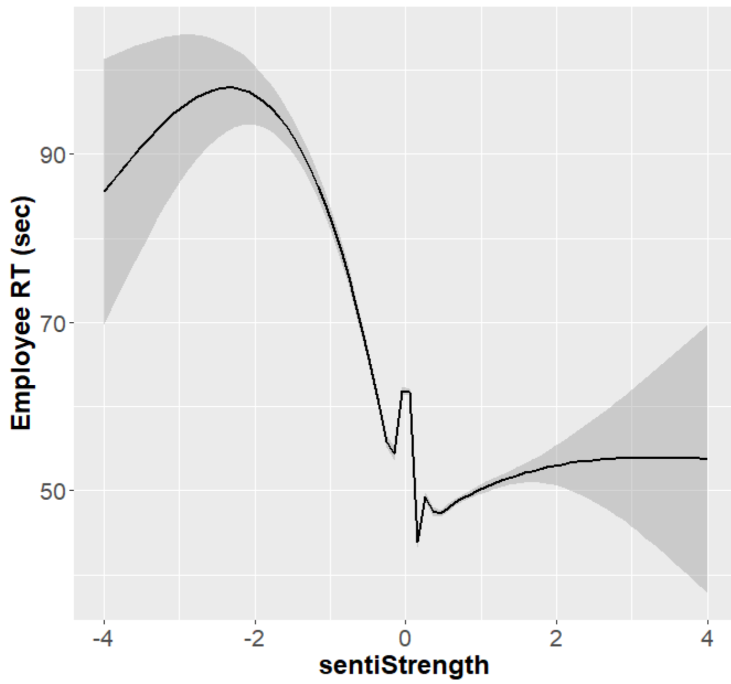


Figure 12. Relationship between emotion in customer messages and employee response time (RT) in subsequent message. (Analysis of 1,447,070 customer messages in 208,210 service conversations)[9].

Figure 13 depicts a pattern that is intriguing on multiple grounds. First, we see that the lowest response time – which is presumably the best agent performance since customers do not like to wait, occurs when customers do not express any emotion. Second, we see that positive customer emotions may seem to add a bit to the overall agent response time, but this is a marginal addition. Third, and perhaps most significantly, the presence of negative customer is seen in Figure 13 to be related to an increase in agent response time. Fourth, the gray areas in Figure 13, which presents the variance (95% confidence interval) around the mean response times represents individual differences among agents in the response times. We see that the variance increases with more extreme positive and negative customer emotions. A possible interpretation is that this variance is due to differences in agent sensitivity to emotion expressed by others, perhaps the extent to which they internalize the emotions of others, or differences in agent emotional intelligence. Clearly this pattern calls for additional research.

Discussion

This chapter describes a new approach for studying customer emotions in service interactions. The approach capitalizes on digital traces that modern day service delivery comprises (Rafaeli et al

[9] Unlike the previous analyses, this analysis is based on emotion detection done by SentiStrength (REF). Values above 0 indicate expression of positive emotion, values below zero are expressions of negative emotion.

2019). To our knowledge ours is a first objective and detailed depiction of the actual emotional encounters that customers express, and that customer-service work comprises. Previous research reported primarily what agents recall, relying on self-report data of agents about their work experience. Since self-reports are subjected to various biases, the analyses of digital traces data that we propose and describe here provide more nuanced and more objective and fine-grained look at the actual nature of service work. The contribution of this chapter is therefore threefold: (a) New methods: we propose automated emotion analysis as a useful tool for research and management of emotion in customer service; (b) Greater breadth: we document dynamics of customer emotions in large-scale samples of real customer-service interactions; (c) Higher granularity: we document emotion dynamics within and throughout service conversations, and within and throughout work shifts of service agents.

The automated sentiment analysis tools that our analyses rely upon can be implemented in real-time and provide insights into customer expressed emotions as a conversation unfolds. Such implementation may offer a way to assess the effectiveness of customer service before a service conversation has ended and without the additional costs and issues associated with customer self-report.

The digital data and newly developed tools for sentiment analyses allow exploration of emotions in large samples of genuine customer-service interactions. The new types of data and methods that we describe and illustrate offer substantial benefits: the research provides objective, unobtrusive views of customer emotions that draw directly from customer expressions, with no self-report intervention and biases (Webb et al. 1966; Xu, Zhang & Zhou, 2019). This chapter thus provides a lens into dynamics of emotions in service that could not be obtained using traditional research methods. The methods and findings we depicted cannot replace traditional research in psychology and organizational behavior, and do not test significance or causality. Rather, the findings unravel new dynamics that should be followed up with more research, both research using traditional experimental methods, and digital traces research that allows inferences of causality. For example, one project we are still working on explores whether customer sentiment influences which customer agents choose to respond to. In this effort we view a response to a customer as a form of attention and thus a reward, which helps us continue to examine the complicated question of whether customer anger is rewarded (Glickson et al, 2019). A second project examines the relationships between customer emotions and agent response times (Altman et al, 2019).

Toward Future Research

Our analyses only touch the surface of types of issues and questions that future research can address following the procedures that we propose. The amount of relevant data, its velocity (the pace with which it is accruing and increasing), and its variety (relevant data comes in many forms, and from many sources both within and outside of the organization) coalesce to a rich and exciting research agenda.

Our findings show that, in contrast to common belief, negative emotions are expressed by only a small proportion of customers (< 10%) and appear in less than 5% of customer sentences. Positive customer emotions are much more common than was previously recognized; positive emotions are also expressed in higher intensity than negative emotions. Some customer conversations include both positive and negative customer emotions (21%), but even this occurs in fewer conversations than conversations with only positive customer emotions (57%). The bottom line is that there is a lot of positive emotion in service delivery, yet we have little research on the triggers, patterns or effects of positive customer emotions. This is therefore a first call of our chapter for future research.

The co-occurrence of positive and negative emotions is shown in our data to be somewhat randomly spread over service agents' work shifts. This suggests that service delivery work is an emotional roller coaster, saturated with affective events (Weiss & Cropanzano, REF). Service agents move back and forth between encounters with customers expressing negative emotion and customers expressing positive emotions. The contrast between encountering positive and negative customer emotions may be responsible the high burnout typical to service agents, rather than only the negative customer emotions. As Sutton & Rafaeli (1991) showed, encounters with contrasting emotions weaken and exhaust people. The co-occurrence and wavering between positive and negative customer emotions may also undelay the tendency of service agents to report a high presence of customer expressions of negative emotions. Contrast effects influence perception and memory (REF), and most available research on the extent of negative emotions relied on self-report methods, which are notoriously susceptible to memory effects. Our second call for future research regards finer testing and verification of these influences.

The evolving emotional footprint within service conversations is an additional new lens our analyses offer; our data show a consistent pattern wherein customer emotions start out negative in almost all conversations, segue into a middle segment which is a plateau with little or no customer expressions of emotions, and an increase toward positive emotion at the end of conversations. Our analyses also suggest that the trajectory of this curve—which varies between customers-- may be meaningful. Our observations of the data plots of large numbers of customers lead us to offer a proposition that the trajectory--the angle of the curve, and / or the point from which customer emotions begin to improve into positive expressions--can predict the extent to which a customer will be content with the service agent, the overall service, and the resolution of the customer service issue *after the conversation has ended*. Our third call for future research is therefore a validation of this prediction (REF – Shelly prediction paper).

Such research is particularly important because it can show the merit of automated emotion assessment conducted in real-time *during a service interaction.* Once the relationship between customer emotions and customer satisfaction is empirically documented, emotion assessments can be exploited not only to predict customer satisfaction, but also to identify and perhaps preempt service failures (Tax and Brown 1998). Current research on service failures relies on customer reports and responses to surveys, which come at a delay, and with limited response rates (cf. Casidy and Shin 2015, Joireman et al. 2013, Smith and Bolton 1998). With further empirical research, post-hoc customer satisfaction might be replaced by real-time monitoring of emotion, which can provide useful tools for managerial decisions or supervisor interventions.

The data, tools and insights shown in this chapter can also provide a practical contribution. The suggested new ways for service organizations to identify customer emotions and can assist service managers in leveraging analyses of customer emotions into their operational procedures. Automated tools can allow a wide range of analyses of all sorts and forms of big-data and connect these data to emotions. Altman et al (2019), for example, use quantitative models of similar data to examine the influences of customer emotions on employee efficiency. Ashtar (2017) used similar data to relate customer emotions to employee's taking of unscheduled micro breaks. Implementations of these and other findings can promote management of operations (George et al. 2016), human resource management (McAbee et al. 2016), and service delivery (Rafaeli et al. 2016, Grewal et al. 2017).

In short, the bottom-line message we wish to convey in this chapter is that a new approach to the study of affect in service offers exciting opportunities. Our data show some novel questions that can be asked, and some novel answers that the Digital Age approach to emotion research can offer. We are certain that our own research only scrapes the tip of a huge iceberg of potential research. We hope our review will stimulate other researchers to join us in this new research adventure.

References

Altman D. et al (2019), MSOM paper under review

Ashtar S. (2017), "The effect of customer emotion and work demands on employee unscheduled breaks: An investigation of chat-based customer service," master thesis, Technion – Israel Institute of Technology. Retrieved from http://www.graduate.technion.ac.il/Theses/Abstracts.asp?Id=30556

Ashtar, Yom Tov, Rafaeli (2019) predicting working paper

Barnes CM, Dang CT, Leavitt K, Guarana CL, Uhlmann EL. Archival Data in Micro-Organizational Research: A Toolkit for Moving to a Broader Set of Topics. *J Manage*. 2018;44(4):1453-1478. doi:10.1177/0149206315604188

Buechel S. and Hahn U. (2017), "EMOBANK: Studying the impact of annotation perspective and representation format on dimensional emotion analysis," In *Eacl 2017,* Vol. 2, 578-585.

George G., Osinga E. C., Lavie D. and Scott B. A. (2016), "Big data and data science methods for management research," *Academy of Management Journal*, 59(5), 1493-1507.

Groth M. and Grandey A. (2012), "From bad to worse: Negative exchange spirals in employee customer service interactions," *Organizational Psychology Review*, 2(3), 208-233.

Jaworek MA, Marek T, Karwowski W. 2018 The scale of Work-Related Affective Feelings (WORAF). 2020;82(September).

Jones NM, Silver RC. (2019) This is not a drill: Anxiety on Twitter following the 2018 Hawaii false missile This Is Not a Drill: Anxiety on Twitter Following the 2018 Hawaii False Missile Alert. 2019;(July). doi:10.1037/amp0000495

Kim S, Mikesell L, Fadem S, Aakhus M. Designing a personalized support tool for patients facing bone marrow transplant. *ACM Int Conf Proceeding Ser*. 2019:41-50. doi:10.1145/3329189.3329206

McAbee S. T., Landis R. S. and Burke M. I. (2016), "Inductive reasoning: The promise of big data," *Human Resource Management Review*, 27(2), 277-290.

Pugh S. D. (2001), "Service with a smile: Emotional contagion in the service encounter," *Academy of Management Journal*, 44(5), 1018-1027.

Qiu L, Lin H, Ramsay J, Yang F. You are what you tweet: Personality expression and perception on Twitter. *J Res Pers*. 2012;46(6):710-718. doi:10.1016/j.jrp.2012.08.008

Rafaeli A, Ashtar S, Altman D. Digital Traces: New Data, Resources, and Tools for Psychological-Science Research. *Curr Dir Psychol Sci*. 2019:096372141986141. doi:10.1177/0963721419861410

Rafaeli A., Erez A., Ravid S., Derfler-Rozin R., Efrat-Treister D. and Scheyer R. (2012), "When customers exhibit verbal aggression, employees pay cognitive costs," *Journal of Applied Psychology*, 97(5), 931-950.

Settanni M, Marengo D. Sharing feelings online: Studying emotional well-being via automated text analysis of Facebook posts. *Front Psychol*. 2015;6(JUL):1-7.

Socher R., Perelygin A., Wu J. Y., Chuang J., Manning C. D., Ng A. Y. and Potts C. (2013), Recursive deep models for semantic compositionality over a emotion treebank. In *Proceedings of the conference on empirical methods in natural language processing (emnlp),* Vol. 1631, Citeseer, 1642.

Strapparava C., and Mihalcea R. (2007), Semeval-2007 task 14: Affective text. In *Proceedings of the 4th international workshop on semantic evaluations*, Association for Computational Linguistics, 70-74.

Thelwall M. (2013), Heart and soul: Emotion strength detection in the social web with sentistrength. In *Proceedings of the CyberEmotions*, 5, 1-14.

Thelwall M., Buckley K., Paltoglou G., Cai D. and Kappas A. (2010), "Emotion strength detection in short informal text," *Journal of the American Society for Information Science and Technology*, 61(12), 2544-2558.

Verhoef, P. C., Antonides, G. and de Hoog, A. N. (2004), "Service encounters as a sequence of events the importance of peak experiences", *Journal of Service Research*, 7(1), 53-64.

Webb E., Campbell D. T. and Schwartz R. D. (1966), *Unobtrusive measures*. Chicago: Rand McNally.

Weiss H. M. and Cropanzano R. (1996), "Affective Events Theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work," *Research in Organizational Behavior*, 18(1), 1-74.

Wirtz J. and Mattila A. S. (2003), "The effects of consumer expertise on evoked set size and service loyalty," *Journal of Services Marketing*, 17(7), 649-665.

Xu H, Zhang N, Zhou L. Validity Concerns in Research Using Organic Data. *J Manage*. 2019;XX(X):1-18.