Adapting Grounded Theory to Construct a Taxonomy of Affect in Collaborative Online Chat

Taylor Jackson Scott, Katie Kuksenok, Daniel Perry
Michael Brooks, Ona Anicello, Cecilia Aragon
University of Washington
423 Sieg Hall
Seattle, WA 98195
{omni at, kuksenok at cs., dbperry at, mjbrooks at, oanicell at, aragon at}uw.edu

ABSTRACT

Distributed collaborative teams increasingly rely on online tools for interaction and communication in both social and task-oriented goals. Measuring and modeling these interactions along different dimensions can help understand, and better design for, distributed collaboration. Affect is one such dimension that can play a crucial role in the dynamics, creativity, and productivity of distributed groups. We contribute an adaptation of the grounded theory methodology as a flexible and extensible means for constructing a taxonomy of affect in text-based online communication. Such a taxonomy can serve as an analytic lens for the continued investigation of the role of affect in creative collaborative endeavors as mediated by communication technology. We describe our modified grounded theory approach and then validate our method by constructing a taxonomy with data from chat logs collected during a longitudinal study of a multi-cultural distributed scientific collaboration.

Categories and Subject Descriptors

K.4.3 [**Computers and Society**]: Organizational impacts – *Computer-supported collaborative work.*

General Terms

Human Factors, Theory

Keywords

Grounded theory, affect, collaboration, taxonomy, text-based communication, computer-mediated communication

1. INTRODUCTION

In a variety of both personal and work settings, communication technologies support distributed groups that rely heavily on text-based forms of communication to achieve their collective goals [3, 5, 10, 15, 20]. As a result of the widespread adoption of these ubiquitous and rapidly changing communications tools, a significant portion of research in related fields has focused on their design and usage.

SIGDOC'12, October 3-5, 2012, Seattle, Washington, USA.

Mehlenbacher's assessment and history of ACM SIGDOC has noted that the field of communication design has expanded to include numerous diverse research areas such as human-computer interaction, computer-mediated communication, interaction design, and collaborative systems. These areas are "united by a common interest in the relationship between text and technology" [18].

The role of communication technologies in creative, distributed collaborations has been of particular interest in creativity research [3, 8, 11]. Recently, Aragon et al. investigated how to foster collaborative creativity in diverse online communities such as scientific researchers and children learning to program [2]. *Affect*–the experience of feeling or emotion–plays a crucial role in these types of creative collaborations, especially influencing communication between group members [3, 10, 13, 15]. As groups work, conversations can range from excitement and confusion to frustration and annoyance, as well as a wide range of other affective states. To investigate this link between affect and the rapidly evolving communication-supporting technologies that shape social dynamics, we have developed a method for constructing a taxonomy of affect to be used as an analytic tool to investigate online text-based communication.

First, we will discuss the increasingly pervasive presence of text-based forms of communication in work practices, as well as the role of affect in the way that this communication is carried out. Second, we will provide an overview of the traditional Strauss and Corbin method for construction of grounded theory that is widely used in the fields of human-computer interaction and design of communication, often in adaptations of grounded methods to alternative analytic ends. This will provide the necessary background for presenting our own adaptation of this method in order to develop a taxonomy of affect in text-based communication. As additional validation of the adapted grounded theory method presented, we will discuss how our taxonomy has been deployed for the coding of chat log data, and how this will support our future research in this area.

2. BACKGROUND

During face-to-face meetings in which the members of a collaborative team are co-located in both space and time, both verbal and nonverbal physical cues play an important role in the way that communication is carried out and processed by the members of the group [14, 15]. However, with the widespread growth in the last few decades of technologies that enable new and increasingly robust modes of remote communication, collaboration is now just as likely to take place between members of a distributed group that do not benefit from the affordances of face-to-face communication [15]. A large portion of this communication takes the form of text, including emails, text messages, and instant messaging chats. These online, text-based forms of communication have become ubiquitous and constitute one of the most important means of contact between members of many distributed groups.

Synchronous online chat differs from other online text-based communication media, such as email, in several significant ways. Since conversations can take place in real time, they capture some of the synchronicity that is associated with face-to-face or voice communication. This synchronicity can greatly enhance the effectiveness and efficiency of this mode of communication by allowing for real-time interaction. However, unlike other real-time modes of communication, text-based chat also has benefits associated with asynchronous communication. Messages can be replied to at the convenience of the correspondents or as dictated by the circumstances of the tasks being performed. Additionally, all of the messages can be logged, providing a persistent record of the conversation. In these regards, this kind of text-based online chat can offer positive aspects of both synchronous and asynchronous forms of communication.

The trace created by using text-based chat communication to mediate creative problem solving can be studied to better understand collaboration. Affect and mood influence creative performance both in individual and collaborative environments [1, 11, 24]. Here, we use Russ's definition of *affect*, 'a feeling or emotion as distinct from cognition' [24]; affect is thus more pervasive than the interrupting neurophysiological experiences of emotions [19]. The expression of affect still plays an important role in these text-based forms of communication, but it takes on forms that are distinct from those found in face-to-face communication [20]. Affect-laden words, emoticons, special abbreviations, deformed spellings, punctuation, and interjections are just a few of the many ways in which the expression of affect has been adapted to text-based forms [13, 15, 16]. These signals, embedded in a detailed trace over time, can help measure the quality and quantity of affect expression in text-based communication between members of a distributed group. The utility of this measurement for further analyses depends on how robust the analytic lens is to the effects that the specific communication medium has on the character of affect expression.

Existing taxonomies of affect and emotion focus primarily on classifying psychophysiological responses to internal states and environmental factors [4, 9, 20]. There are several conflicting theoretical models of affect and emotion in multiple fields. Examples include the dimensional models of Russell [25], the emotion wheel of Plutchik [20], and the distinction between basic and complex emotions [4, 9]. Emotion is often measured via facial expression, vocal features, and body posture as the physical *expression* of emotion (rather than focusing on internal *state*) [28].

Researchers interested in understanding how we express and perceive emotions rely primarily on analyzing these physical forms of expression [28]. This is not possible when attempting to measure the expression of affect in text-based communication from a chat log; furthermore, the affordances of the medium can lead to individuals adjusting communication practices to express affect via text in ways that they would not do using the spoken word and thus are not well-accounted for in existing taxonomies. Our work bridges the gap between research that considers the measurement of affect, and the design of communication media that can fundamentally shape human expression.

3. OVERVIEW OF GROUNDED THEORY

The grounded theory (GT) method is traditionally described as "a qualitative research method that uses a systematic set of procedures to develop an inductively derived grounded theory about a phenomenon" [29]. The goal of this method is to generate a theory that emerges from the data being comparatively analyzed, rather than the application of an existing theory to answer a research question [12, 29]. The method is especially well suited to producing theories of interactions between different social units, and is widely used in many social science fields [6].

In order to analyze the data and build a theory that is grounded in it, Strauss and Corbin suggest three types of coding activities – open, axial, and selective [29]. While they are generally carried out in sequence, they are also often used iteratively as the research progresses, taking advantage of the emergent and reflexive properties of this method [6]. These procedures form the core of this methodological approach and are the main processes by which the data is used to generate a theoretical framework.

3.1 Open Coding

During open coding, text data, such as field notes or interview transcripts, is examined line by line, the main concepts and categories are identified, and their properties and dimensions are initially captured through the use of memos that discuss the researchers' ideas behind the codes. The concepts captured in these memos can be seen as the core units of the theory being developed. The similarities and differences between data points are examined, then named and recorded in the memos. This phase of coding is an open process during which all pieces of data are of interest to the researcher, and few if any restrictions are placed on what data gets coded and how it is conceptualized.

3.2 Axial Coding

During axial coding, categories, concepts, and codes are related to one another by linking them around the axis of a single category at the level of their properties and dimensions. In order to understand how these categories and codes relate to one another, Corbin and Strauss suggest the use of a "paradigm model" that takes into consideration the relationships between conditions, context, actions/interactions, and consequences. The basic idea of this model is to systematically propose linkages between these aspects and then look back to the data for validation. This paradigmatic model is then used to link sub-categories with their respective categories in a way that reveals an underlying structure of the codes produced during open coding. Generally, axial coding proceeds until a level of "theoretical saturation" is reached whereby gathering or examining new data does not lead to the emergence of substantially new structure.

3.3 Selective Coding

The first major goal of selective coding is the formulation of a core variable or category to which all other categories and codes can be related. At this point, open and axial coding processes cease, and only those categories and variables that can be related to this core variable continue to be coded as the formulation of the theory proceeds. Strauss and Corbin point out that this core category should be able to explain variation as well as contradictory evidence found in the data. The core variable represents a type of narrative that is grounded in the data by which the categories identified during axial coding are linked.

Finally, as a means of refining the theory produced through the selective coding process, Strauss and Corbin suggest that the researcher review the theory to check for internal validity and logic; attempt to account for underdeveloped categories; eliminate any excess categories; and to validate the theory (as might be accomplished through a high-level comparison with the original data).

3.4 Grounded Theory in Context

Qualitative data analysis methods grounded in data enable the discovery of emergent themes, rather than focus empirical investigation on pre-specified hypotheses. These methods have been critical in constructing and refining theories of social phenomena, including in human-computer interaction topics. Schoonewille et al. used GT to better model and understand developer comprehension of software documentation and then validated their model using a cognitive theory of multimedia [27]. Power and Moynihan used an adapted GT approach in constructing a framework to explain the situational variety of styles of requirement documentation, as well as a three-part scheme for classifying these requirements that was a direct result of their GT coding [22]. Finally, Razavi and Iverson produced a theory of end-user information sharing behavior in a personal learning space using grounded theory methods that were enacted in a similar fashion to our own [23].

In research of human interaction with information systems, grounded theory methods are used not only for construction of a theory, or as part of a mixed-method case-study approach, but also as a means to refine an initially hypothesized theory [17, 26]. In this case, the initial theory can be modified, refined, or further informed by themes that emerge from the application of open, axial, or selective coding to qualitative data. Our adaptation of this approach, on the other hand, takes the route of treating GT as an intermediate analytic step that results in a taxonomy to be used in subsequent analyses. Our expectations for this taxonomy extend beyond the construction of a theory, including, for instance, the need to automatically detect instances of codes in a large-scale dataset and perform statistical analyses. The goal of using this taxonomy for a purpose not typically part of the GT method led to the adaptations we propose.

The primary characteristic of the analytic processes of GT is closeness to the data. In studying affect in text-based communication, nuanced means of communicating and expressing affect are of key interest to research, and pose challenges to existing taxonomies of affect which are not amenable to the peculiarities of text-based, distributed expression [4, 9, 20, 28]. Not only do these taxonomies rely on implicit physical characteristics not present in text for classification, they are also not ideal for capturing subtle affective states such as confusion or agreement. These other forms of affect are just as important as classical categories of emotion when attempting to account for all of the factors contributing to the dynamics of the group. The capacity to systematically extract previously unacknowledged themes is inherent in the GT methodology and is crucial in this task. Nevertheless, the direct application of GT traditionally results in a theory or model of the data, which is not the purpose for which we want to use these methodologies.

4. CONSTRUCTING OUR TAXONOMY

Our dataset is comprised of four years of chat logs created by the cross-cultural collaboration among members of the Nearby Supernova Factory (SNfactory), an astrophysics collaboration of approximately 30 core members; about half of the scientists are located in the U.S. and the other half in France. These scientists are studying Type Ia supernovae, a specific type of stellar explosions that have a consistent brightness, allowing their distances to be effectively measured over time and thus trace the expansion history of the universe. The group operates their telescope remotely three nights per week; during such operation, numerous decisions must be made quickly and collaboratively despite the fact that many of the team members have never met each other and come from differing cultural backgrounds. Chat is the team's primary means of communication during telescope operation [2].

These situational factors shape the team members' expression of affect as they carry out their work and communicate with each other. Over the four-year span of our chat log, conversations range over excitement at new findings, frustration with faulty software or hardware, confusion with incoming data, and many other affective states. There are a total of 485,045 chat messages, many produced by automated programs ("bots") using the chat protocol to relay changes in the state of the world (sunrise/sunset; weather; telescope settings, etc.). One of the primary concerns when developing our taxonomy was to account for the nuanced and specific ways that affect is communicated and expressed in a text-based medium. We found that many existing taxonomies were created to characterize affect predicated on implicit physical representations such as facial expressions or tonal inflection [4, 9, 20, 28]. This stands in stark contrast to much of the expression of affect in text-based communication which relies heavily on explicit statements of emotion and text features such as emoticons and punctuation [13, 20]. Our application of GT was specifically intended to develop and refine a taxonomy that captured these types of affective expressions. Although this construction of a taxonomy for coding is not a typical use of GT, an appreciation for the method's closeness to the data as well as the method's ability to identify and group themes made it an ideal candidate for adaptation to our needs. Whereas the codes generated and applied during grounded theory are generally used to provide structure and inform the development of a theory, we were specifically refining these codes into a taxonomy that could be used as a coding scheme in its own right for the further analysis of our data.

During traditional open coding, data is initially organized into concepts and themes [29]. Using this approach, we explored the data in an unrestricted manner through a careful line-by-line reading of portions of the chat logs. Given that this stage of the GT method is specifically geared towards openness, the need for adaptation was minimal. We initially coded anything and everything that was of interest, not just affect, but also accounted for instances of creativity, collaboration, and other events significant to the group. Due to the scale of our data, it was not possible to perform open coding on the entirety of the data set, so we strategically sampled areas that contained high volumes of interaction between the participants in order to maximize our chances of finding significant and interesting phenomena.

Axial coding enabled us to focus the scope of what we would account for in our taxonomy. In addition to this substantive dimension, we also began to explore the inclusion of two other separate axes. As we related the instances of affective expression to one another, we found it useful to note their intensity (high or low) as well as their valence (positive, neutral, or negative) as is commonly done in sentiment analysis. These measures provided an additional fine-grained characterization of our substantive codes as well as a means to resolve ambiguities when applying codes that were not explicitly positive or negative depending on the context. Along with the substantive axis, the intensity and valence axes formed the overall paradigmatic model by which the codes produced during open coding were grouped and refined around a central set of codes.

Selective coding involves the formulation of a core variable or category to which all other categories and codes can be related to in a sufficiently significant way to be considered a substantial part of the final theory being developed [29]. For us, this core variable took the form of an 'affect' category to which all of our codes in the taxonomy were being related. One primary difference between the traditional GT approach and ours was that we still continued an iterative approach to open and axial coding during our selective coding. This was done because through our selective coding process, we continued to encounter new and significant points of interest in our data. As these new codes were identified and defined, they were combined into the selective coding process as part of our attempt to reach the theoretical saturation that generally signifies the completion of this phase in traditional GT [12]. This selective coding process was also an opportunity to explore the internal validity of our taxonomy by evaluating how well it accounted for affective expressions in our data.

4.1 Plutchik's Wheel of Emotion

After selectively coding our data, we were left with a core category (affect) to which we had related all of our other variables in order to form a theoretical framework that could account for affective expression encountered during open coding. We had also begun the process of validating, expanding, and trimming our codes through the iterative application of our categories to the data in order to check the internal validity of our coding scheme and the resulting taxonomy. At this point in the grounded theory process, it is generally accepted that relevant literature will be reviewed in order to better situate the emerging theoretical model within the existing body of research in that area [12]. This step in the grounded theory method was well suited to the creation of a taxonomy of affect without replacing or reinventing existing taxonomies. We hoped to account for the shortcomings of taxonomies that were not specifically capable of addressing the variations of affect expression that are present in text-based communication mediated by online chat.

Through our review of existing taxonomies of affect and emotion, we found that Plutchik's Wheel of Emotion was closely aligned with what we had been seeing and coding for in our data. Although we have been careful to make a distinction between affect and emotion, they are still closely related. In fact, not only were there numerous codes in our taxonomy that were not present in the Plutchik wheel, there are several emotions from Plutchik that were not in the taxonomy we had created but were still found to be applicable to our own data. Additionally, the inclusion of the Plutchik emotions also ensures that our own work builds on and extends existing theories of affect and emotion.

Ultimately, the application of this adapted grounded theory approach was our solution to the problem of attempting to translate a very large body of work on affect and emotion into a more appropriate and useful analytic lens. This lens can then be used to examine the specific types of affective expression present in our data because it accounts for the distinct ways in which these expressions are molded by the text-based medium.

5. RESULTS

5.1 Taxonomy in Use

During several months, five members of our research team iterated on developing a coding scheme as part of the adapted grounded process we have described. The resulting taxonomy includes substantive codes reflecting affect state expression (listed in Figure 2 below), as well as valence codes relating to how

time	*speaker*	*message*
05:58:41	Alice	ok, so where was the f***ing SN on the image? #1: interest / anger / high / negative #2: annoyance / confusion / low / negative #3: interest / frustration / high / negative
05:58:55	Alice	was it the bright blob? #1: interest / anger / high / negative #2: considering / low / negative #3: interest / neutral
05:59:03	Ben	5876 absorption is much wider than the H alpha in v space #1, #2, #3: no affect
05:59:18	Ben	Oh hmmm. #1, #2, #3: considering / neutral
05:59:28	Ben	Lemme see what [the] coordinates were... #1, #2, #3: no affect
06:13:07	Charlie	is it "well-developed"? #1: interest / neutral
06:13:18	Alice	Should be an interesting experiment. #1, #2: anticipation / low / positive #3: interest / neutral
06:13:19	Dana	yes #1, #3: agreement / neutral #2: no affect
06:13:20	Dana	big!! #1: excitement / agreement / high / positive #2, #3: excitement / low / positive

Figure 1. Two examples of conversations from our dataset, with anonymized speaker names and excluding any identifying detail. Each segment was coded by three members of the research team; their annotations are shown below each line.

positive, *negative,* or *neutral* a message is overall, and intensity codes for *low* or *high* expression intensity (where 'neutral' valence does not call for an intensity code). Substantive codes are not mutually exclusive, and can be combined with valence and intensity labels for greater flexibility. For example, a sarcastic comment can express nuanced affect in this context, including "frustration/negative/high" during particularly stressful periods, and "amusement/positive/low" during less demanding times (see Figure 1 above for examples).

Messages could also be coded as "no affect" to systematically distinguish messages that had been coded and identified as expressing no identifiable affective state, and those which were yet to be coded. We coded approximately 5% of the total chat log data with the final taxonomy, utilizing a team of three primary coders and five additional coders, all part of the research team. For several weeks, coders focused on applying substantive codes and 'no affect,' and then the additional intensity and valence axes were added to the affective coding scheme. A summary of how many messages were coded, and by how many coders, is shown in Table 1. Of 35,614 messages coded, 15,942 (45%) were coded as 'no affect' by at least one person – although, as the second example in Figure 1 shows, 'no affect' can be plausibly incident with more neutral affect codes, depending on interpretation.

In many cases, multiple substantive codes, those codes which capture the nature and meaning of a message, may apply, such as *annoyance* and *frustration* applying to these three messages, sent by the same person: "Did I see a bunch of = vs == in there??? / WHAT / WHO DID THAT". The theoretic basis of the Plutchik taxonomy includes relationships between codes that are more or

Less Intense	→	More Intense
pride serenity	amusement joy	ecstasy
agreement acceptance	supportive trust	gratitude admiration
tired distraction	disbelief surprise	amazement
considering interest	relief anticipation	excitement vigilance
apologetic pensiveness	embarrassment sadness	grief
apathy boredom	disgust	frustration
disagreement apprehension	confusion fear	terror
annoyance	impatience anger	rage

Figure 2. Our taxonomy, arranged from less to more (left to right) intense affect expression, and with more typically positively-charged instances at the top and more typically negatively-charged at the bottom, with many in the middle, such as 'interest,' capable of being expressed in a positive as well as negative context.

less intense variants of one another (e.g. apprehension/fear/terror). Additionally, multiple substantive codes can be used simultaneously, such as "anger / confusion / low / negative" expressed in a conversation about error-prone software. Example messages where this expressiveness is especially useful are included in Figure 1. Of the messages coded, 1,599 were coded with multiple substantive codes simultaneously by at least one coder (129 were coded with more than two substantive codes by at least one coder).

This process took several months, and leveraged a tool for coding chat logs that was developed within our team (shown in Figure 3). The tool was developed over the course of a year, simultaneously with the creation of the taxonomy, to make the coding of chat logs faster and easier, and coded data more accessible for analysis (via storage in a central relational database and a carefully designed user interface). We plan to release this tool to the public in the near future.

Table 1. Summary of coding progress using the included taxonomy. This table shows the number of messages coded with substantive affect codes only, the number of messages coded along all three axes, and the distributions of coding across different numbers of independent coders.

#coders	#messages coded with substantive codes or 'no affect'	#messages coded with substantive codes, valence, and intensity
1	18,843	8,399
2	5,274	1,073
3	2,704	403
>3	537	17
all	**27,344 (5.64%)**	**9,892 (2%)**

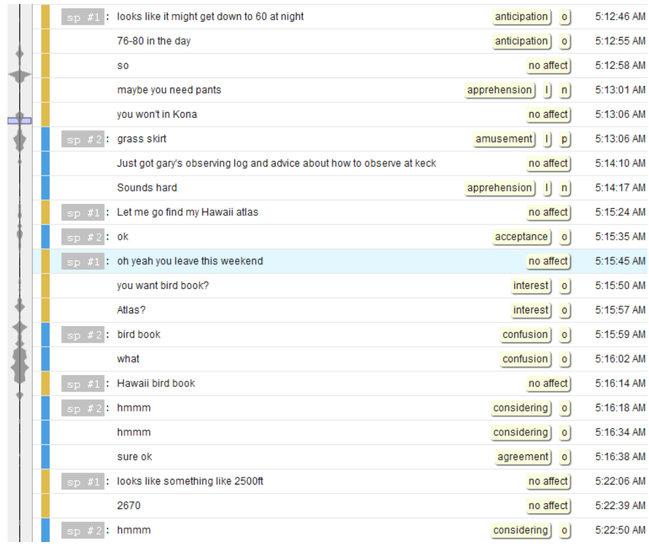

Figure 3. A screenshot of the coding tool, developed by the team, as seen by an individual coder applying codes. ("o" is neutral, "l" and "h" are low and high intensity, and "n" and "p" indicate negative and positive valence)

5.2 Inter-Rater Reliability

Although the grounded methodology itself does not always call for the calculation of inter-rater reliability, verifying that human annotators can reliably apply codes validates the use of a grounded taxonomy for further analytic steps. However, the construction of this taxonomy does not necessarily produce one that is non-exclusive. Applying multiple codes at once, while necessary for capturing nuanced dimensions of affect expression, for example, violates one typical assumption of reliability metrics: the exclusivity of codes.

Cohen's kappa [7] and other widely used reliability metrics that calculate the reliability based on *observed agreement* to *chance agreement* tend to assume exclusive application of codes. The main problem with such an application is that when coders can apply multiple codes per item, the standard estimate of the probability of chance agreement becomes erroneous (underestimating the true probability). We analyzed our coded data using a modified version of the kappa statistic, which overcomes the problem of non-exclusive code application through a Monte Carlo simulation.

We first extended the *observed agreement* term to work for non-exclusive codes. We defined agreement about a particular code on a single chat message in the following way: if more than half of the people who coded the message said that the code was present, then they agree. If all of them said that the code was absent, then they also agree. Any other combination is disagreement. For comparison, traditional kappa calculations, which deal with only two coders, also consider the coders to be in agreement when they chose the same rating. We used this definition of agreement because our coders had many non-exclusive codes available, and because *not* using a code on a specific message does not strongly imply that the coder disagrees with that code, only that it was less appropriate than other codes available (e.g. *rage*, *annoyance,* and *fear* may co-occur).

Estimating the probability of coders agreeing by chance is more complex. Since chat messages may have variable numbers of coders, and coders may choose to apply variable numbers of codes, it is difficult to compute the probability of chance agreement directly. We developed an estimate of the probability of chance agreement based on a Monte Carlo method. We first calculate the probability of choosing each code for each individual coder, and the probability of applying specific numbers of codes for each individual coder. This gives us a profile of each coder's general behavior, independent of which chat message is being coded.

Next, we randomly simulate ratings for a very large number of messages. For each message we are simulating, we decide which coders are going to rate it, based on the proportion of messages rated by those coders in the dataset. Then, for each of those coders, we randomly choose a number of codes to apply, sampling from the distribution we already calculated for that coder. Each of these codes is randomly selected from the coder's prior distribution of code choices. Counting the number of these simulated messages where agreement occurred allows us to estimate the probability of random agreement. This approximates the typical measure of chance agreement in the two-rater exclusive-code case, but generalizes to our more heterogeneous data. The Monte Carlo simulation continues until all probability estimates are stable to within 0.0001, generally requiring about 2 million messages to be simulated.

To finally calculate this modified kappa, we divide the difference between the rates of observed and chance agreement by the difference between one and the rate of chance agreement [7]. Some example reliability measures over our own data are shown below in Table 2.

Table 2. The observed and chance agreements, and resulting modified-kappa reliability scores, for 6 of the most common codes in our taxonomy over a 35k message dataset

Code	Obs. % Agreement	Prob. Chance Agreement	Kappa
interest	0.925	0.609	0.808
amusement	0.933	0.827	0.611
considering	0.931	0.864	0.490
agreement	0.954	0.909	0.491
confusion	0.906	0.755	0.615
annoyance	0.929	0.693	0.770

6. DISCUSSION

The method that we have applied to this specific problem area and data set is based on embracing the qualitative process. The validation or evaluation of this method poses several key limitations that we hope to address. We also frame the work that we have completed thus far in the context of a larger research agenda, with the possibility of extending our method to account for other data sets and phenomena not specific to affect.

6.1 Limitations

Given the size of the data set that we are coding (485,045 messages), it is reasonable to assume that there could very well be unique instances of affective expression that we have not yet encountered during the grounded theory coding process itself, or the subsequent coding of our data with the resulting taxonomy. Although we have tried to anticipate and account for this not only through rigorous sampling of the data, but also through the integration of the Plutchik taxonomy with our own in order to make it more robust and flexible; we acknowledge that there may be affect in the data that is not specifically accounted for in our taxonomy. Therefore, we do not claim that our taxonomy is exhaustive, but only that it has thus far successfully accounted for the affective content we have encountered.

In both formulation and application, the coding scheme is not comprised of mutually exclusive codes. While it is often the case when creating or using a coding scheme to have only one specific code that is applicable to any given piece of data, we found this restriction too limiting in effectively capturing the variety and subtlety of the affective content that we sought to identify in our data. The flexibility of combining concepts afforded by non-mutually-exclusive coding, such as in open coding, is in tension with analytically-motivated exclusive coding along each dimension, typical to axial coding. The decision to favor the flexibility of non-mutually-exclusive codes influences the interpretation of coded data. For instance, for a given line of chat, if code A is applied, but not code B, code B might still have been justifiable, but subjectively less than A. If one coder applies A and B, and another only A, it is a different sort of disagreement than if one of the coders only applied code C. There are also consequences for the measurement of inter-rater reliability. Standard formulations of reliability metrics are not strictly applicable in this case; a modified kappa methodology appropriate here was detailed in section 5.2.

Despite the difficulties introduced by foregoing exclusivity in coding, this decision grants the taxonomy more expressive power. There are instances when a single line might share two or more codes (such as anger, frustration, and annoyance all occurring simultaneously). Because these affective states often co-occur, we find it valuable from an analytical standpoint to retain access to this granularity of coding. We deliberately chose to avoid flattening the instances of codes occurring in our data by collapsing several co-occurring codes into a single unified code. It was more practical to allow the co-incidence of codes, such as anger and frustration, or surprise and frustration, rather than increasing the taxonomy, potentially combinatorially, in response to the complexities of affect expression in chat.

Finally, we constructed this taxonomy to answer specific questions about the role of affective expression in the dynamics of a particular distributed collaborative team (SNfactory). Thus far, we have not attempted to apply it to another corpus of chat logs or other forms of text-based communication. We plan to address these limitations through our future research which will explore the usefulness of this taxonomy for other chat data sets.

6.2 Future Work

The method described here resulted in the construction of a robust taxonomy of affect that is firmly grounded in our data set and builds on a large body of related work. Unlike traditional grounded theory approaches, our method focuses on using aspects of the methodology that maintain closeness to emergent themes in the data to construct an analytic lens both sufficient for our data and flexible enough to be used in other types of investigation. As Charmaz notes, even finished grounded theories are somewhat open-ended and the constructions of concepts are able to shape both the process and the final product [6]. We expect that this will be reflected in the ongoing construction and refinement of this taxonomy as we continue to apply it to our data; not only through our own reflexive engagement with our taxonomy, but also through the dialectic relationship between the taxonomy and the data it was derived from.

We intend to further validate our approach by utilizing the taxonomy to code other data sets for affective content. This would provide the opportunity to see how well it captures and accounts for this property more generally. We can then compare inter-rater reliability scores to address performance between the two corpora. It would be particularly interesting to investigate how appropriate this taxonomy is for a wider variety of text corpora. Additionally, since affect is only one example of the phenomena that play an important role in collaborations, it would also be appropriate to apply the same methodological approach to the construction of a taxonomy for some property other than affect. This could be achieved through a reiteration of the adapted grounded theory approach to our own data set, or to some other unique data set.

Understanding affect in collaborative work is an important topic that can reveal much about how communication is conducted via these types of real-world exchanges, as well as how we might design new technologies to support them. This work can also further the development of a model of how collaborative communication takes place and how it impacts group dynamics. The development of this adapted grounded theory approach and the construction of our taxonomy are not ends in and of themselves, but are rather first steps in a more comprehensive understanding of how the affective states of participants in distributed collaborations are related to the dynamics and productivity of these teams.

7. CONCLUSION

Our ongoing research on the role that affect plays in distributed group collaborations motivated the development of a taxonomy that accounts for the distinctive expression of affect that takes place in text-based online communication. We drew upon existing bodies of research on both emotion and computer-mediated communication to inform our approach, and ultimately used a novel adaptation of the grounded theory method to construct an appropriate taxonomy. We wanted to account for the nuanced and specific ways that affect is communicated and expressed in a text-based medium, and existing taxonomies of emotion were found to not be a good fit for this goal. This adapted grounded theory approach was our solution to the problem of attempting to translate a very large body of work on affect and emotion into a more appropriate and useful analytic lens that accurately reflected the phenomena of affect found in our data. The resulting taxonomy has been used to code a large corpus of chat logs collected during a longitudinal study of a distributed scientific collaboration. We hope that other researchers in this area will find both the method and the taxonomy we have presented to be of use in their own studies.

8. ACKNOWLEDGMENTS

We deeply appreciate the reviewers' careful reading of our work which greatly strengthened this paper, and we thank the scientists of the SNfactory collaboration. This work was funded in part by an NSF Graduate Research Fellowship in Computer Science.

9. REFERENCES

[1] Amabile, T.M., Barsade, S.G., Mueller, J.S., and Staw, B.M. 2005. Affect and Creativity at Work. *Administrative Science Quarterly 50*, 3 (September 1, 2005). 367-403.

[2] Aragon, C. R., Poon, S., et al. 2009. A tale of two online communities: fostering collaboration and creativity in scientists and children. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition* (Berkeley, California, USA, 2009). C&C '09. ACM, New York, NY, 9-18.

[3] Aragon, C.R. and Williams, A. 2011. Collaborative creativity: a complex systems model with distributed affect. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada, 2011). CHI '11. ACM, New York, NY, 1875-1884.

[4] Barrett, L.F. 2006. Solving the Emotion Paradox: Categorization and the Experience of Emotion. *Personality and Social Psychology Review 10*, 1 (February 1, 2006). 20-46.

[5] Bullinger, A.C., Hallerstede, S.H., et al. 2010. Towards research collaboration – a taxonomy of social research network sites. In *Proceedings of the Sixteenth Americas Conference on Information Systems* (Lima, Peru, August, 2010). AMCIS '10. 12-15.

[6] Charmaz, K. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*. Sage Publications.

[7] Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement 20*, 1 (April 1, 1960). 37-46.

[8] Corbin, J.M. and Strauss, A.L. 2008. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.

[9] Csikszentmihalyi, M., 1997. *Creativity: Flow and the psychology of discovery and invention*. New York: Harper Collins Publishers.

[10] De Silva, L.C., Miyasato, T., and Nakatsu, R. 1997. Facial emotion recognition using multi-modal information. In *Proceedings of 1997 International Conference on Information, Communications and Signal Processing.* ICICS '97. 397-401.

[11] Ekman, P. 1992. An argument for basic emotions. *Cognition & Emotion 6*, 3-4 (1992/05/01), 169-200.

[12] Fitzpatrick, G., Kaplan, S., and Mansfield, T. 1996. Physical spaces, virtual places and social worlds: a study of work in the virtual. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work* (Boston, Massachusetts, USA, 1996). CSCW '96. ACM, New York, NY, 334-343.

[13] George, J.M. and Zhou, J. 2002. Understanding when bad moods foster creativity and good ones don't: The role of context and clarity of feelings. *Journal of Applied Psychology 87*, 4, 687-697.

[14] Glaser, B.G. and Strauss, A.L. 1967. *The discovery of grounded theory; strategies for qualitative research*. Aldine Pub. Co., Chicago.

[15] Hancock, J.T., Landrigan, C., and Silver, C. 2007. Expressing emotion in text-based communication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA, 2007). CHI '07. ACM, New York, NY, 929-932.

[16] John-Steiner, V. 2000. *Creative collaboration*. Oxford: Oxford University Press.

[17] Kraut, R., Egido, C., and Galegher, J. 1988. Patterns of contact and communication in scientific research collaboration. In *Proceedings of the 1988 Conference on Computer-supported Cooperative Work* (Portland, Oregon, USA, 1988). CSCW '88. ACM, New York, NY, 1-12.

[18] Liu, H., Lieberman, H., and Selker, T. 2003. A model of textual affect sensing using real-world knowledge. In *Proceedings of the 8th International Conference on Intelligent Sser Interfaces* (Miami, Florida, USA, 2003). IUI ’03. ACM, New York, NY, 125-132.

[19] Matavire, R. and Brown, I. 2008. Investigating the use of "Grounded Theory" in information systems research. In *Proceedings of the 2008 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT research in Developing Countries: Riding the Wave of Technology* (Wilderness, South Africa, 2008). SAICSIT '08. ACM, New York, NY, 139-147.

[20] Mehlenbacher, B. 2011. The evolution of communication design: a brief history of the ACM SIGDOC. In *Proceedings of the 29th ACM International Conference on Design of Communication* (Pisa, Italy, 2011). SIGDOC ’11. ACM, New York, NY, 249-256.

[21] Moore, B.S. and Isen, A.M. 1990. *Affect and social behavior*. Cambridge [England]: Cambridge University Press.

[22] O'Neill, J. and Martin, D. 2003. Text chat in action. In *Proceedings of the 2003 international ACM SIGGROUP Conference on Supporting Group Work* (Sanibel Island, Florida, USA, 2003). SIGGROUP ’03. ACM, New York, NY, 40-49.

[23] Plutchik, R. 1980. A general psychoevolutionary theory of emotion. In *Emotion: Theory, Research, and Experience: Vol. 1. Theories of Emotion*, R. Plutchik and H. Kellerman Eds. Academic press, 3-33.

[24] Power, N. and Moynihan, T. 2003. A theory of requirements documentation situated in practice. In *Proceedings of the 21st Annual International Conference on Documentation* (San Francisco, CA, USA, 2003). SIGDOC ’03. ACM, New York, NY, 86-92.

[25] Razavi, M.N. and Iverson, L. 2006. A grounded theory of information sharing behavior in a personal learning space. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work* (Banff, Alberta, Canada, 2006). CSCW ’06. ACM, New York, NY, 459-468.

[26] Russ, S.W. 1993. *Affect and creativity: the role of affect and play in the creative process*. L. Erlbaum Associates, Hillsdale, N.J.

[27] Russell, J.A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology 39*, 6, 1161-1178.

[28] Sarker, S., Lau, F., and Sahay, S. 2000. Using an adapted grounded theory approach for inductive theory building about virtual team development. *SIGMIS Database 32*, 1, 38-56.

[29] Schoonewille, H.H., Heijstek, W., et al. 2011. A cognitive perspective on developer comprehension of software design documentation. In *Proceedings of the 29th ACM International Conference on Design of Communication* (Pisa, Italy, 2011). SIGDOC ’11. ACM, New York, NY, 211-218.

[30] Zhou, L. and Dongsong, Z. 2005. A heuristic approach to establishing punctuation convention in instant messaging. *Professional Communication, IEEE Transactions on 48*, 4, 391-400.