# Reading between the lines

## Detecting emotion-abstract language use in the web-based treatment “Look at your drinking”

Marijana Krstić

Faculty of Behavioural, Management and Social Sciences

University of Twente

Supervisors

Wouter Smink, MSc

dr. Anneke Sools

dr. Erik Tjong Kim Sang

In cooperation with

Abstract

Language is an important mediator of psychotherapeutic processes. The proportion emotion-abstract language is reported as a valuable factor in achieving positive outcomes in therapy. In contemporary times, the context of psychotherapy is moving to virtual surroundings. The most promising results of web-based treatments so far are found in domain of addiction interventions. The tools LIWC and DAAP are text mining techniques that can identify emotion-abstraction related categories and Referential Activity (consisting of Arousal, Symbolizing, and Reorganizing Phases) in text. This study demonstrates the potential of using these techniques in detecting emotion-abstract language and importance of further analysis of its use in web-based treatments. In this study, existing LIWC and DAAP programmes were applied to a sample of 9 clients involved in web-based intervention "Look at your drinking", focused on excessive drinkers. Mixed methods approach enabled analysis of the obtained results within context, by allowing the researcher to see how distribution of LIWC categories or RA values was manifested in the emails of the participants. Results of the study show that in in web-based treatment for excessive drinkers social, cognitive, and affective processes are the most prevalent categories. Referential activity is shown to be measurable in online surrounding, indicating how increased Referential Activity affected engagement of both clients and counsellors. Referential Activity emerged in cycles during the treatment, as identified in face-to-face treatments. Both LIWC and DAAP scores reach higher values within the same timeframe. In the light of collected data, emotion-abstract language shows promising results for further development and improvement of delivering web-based interventions.

*Keywords*: emotion-abstract language, text mining, LIWC, DAAP, online therapy

Reading between the lines

Detecting emotion-abstract language use in the web-based treatment “Look at your drinking”

## Introduction

Throughout the history, language has shown to be an important factor for understanding the human mind. “Language is the most common and reliable way for people to translate their internal thoughts and emotions into a form that others can understand” (Tausczik & Pennebaker, 2010, p. 25). The opportunity to experience this form of interpersonal exchange has set a cornerstone for various psychotherapeutic approaches. Language, as an important mediator of such processes, encompassed Freud’s “talking cure” therapy, which acknowledged the impact of conversation with the aim of reaching out for the underlying mechanisms in person’s understanding. Simultaneously, lexical hypothesis underpinned that personality traits are reflected in the language itself. Narrative psychology perceives story-like formats as relevant components of making sense of one’s lives (Pennebaker & Seagal, 1999), making the structure of language and structure of thought intertwined (Bruner, 1991). Independent from the preferred approach, psychotherapy entails identification and exchange of cognition, emotions, beliefs, thinking patterns and comprehensions of relationships. According to the American Psychological Association, psychotherapy is defined as a cooperative treatment that is grounded in dialogue, enabling a supportive relationship between the client and the therapist (American Psychological Association [APA], n.d.). Knowing that language and spoken words are undoubtedly crucial aspects of human interaction, the question that may arise is: How does language, as an essential component, find its place in modern era?

Due to the fact that the core of psychotherapeutic process has not fundamentally changed in the last 70 years (Imel, Steyvers, & Atkins, 2016), one of the main focuses of this field is how

to utilize its well-established procedures into rapidly changing world. Nowadays, we are witnessing inevitable integration of modern technologies into our everyday lives, which led to many conceptual changes in terms of communication, amount of data traffic, and accessibility of health services. We are living in a society in which human activity is vastly and growingly mediated via digital devices, leading to aggregation of "big social data" (Lambiotte & Kosinski, 2014). Big social data consists of digital footprints, which are either freely available online materials, such as social media posts, reviews, and browsing histories, or data collected from companies and institutions for different purposes (Kosinski, Wang, Lakkaraju, & Leskovec, 2016). So far, analysis of such data has been employed for multiple purposes, from predicting psychological traits, online marketing, to personalized search engines and advertisements (Lambiotte & Kosinski, 2014). Machine learning methods that can tackle this type of data are now being integrated into the field of psychotherapy (Imel et al., 2016; Murphy, 2015; Tanana et al., 2015).

Since the written or spoken word is no longer bound to spatiotemporal context, researchers have the opportunity not only to access great amount of transcribed material or written correspondence, but also to endow the field of psychotherapy with newly developed analysis techniques. The utilization of computer programmes that tap into the linguistic nature of psychotherapy has already begun 30 years ago (McCarthy et al., 2011), and it is evolving fast. This rapid expansion allows gaining deeper understanding of the online psychotherapeutic content and establishing the grounds for new field of web-based psychotherapy.

**Web-based psychotherapy**

Psychotherapy consists of active communication about feelings, information and relationships between the therapist and the client (Canfield, Walker, & Brown, 1991). Nowadays,

using telehealth or telepsychology[1] has become increasingly prevalent in practice (Rees & Stone, 2005). The expansion of fields of application of this form of delivery of psychotherapeutic services has already reached mental health (Postel, Haan, Jong, & Ph, 2008), including depression (Clarke et al., 2005), anxiety (Richards, Richardson, Timulak, & Mcelvaney, 2015), addictions (Rodda, Booth, Vacaru, Knaebe, & Hodgins, 2018; Strecher, Shiffman, & West, 2005), and chronic diseases (Andry et al., 2008; Sinha, Porter, Hons, & Wilson, 2018).

According to Andersson and colleagues (as cited in Postel, De Haan, Ter Huurne, Becker, & De Jong, 2010), internet interventions can be differentiated based on the distribution of involvement of the client and therapist in the following manner: (1) pure self-help, (2) self-help with therapist assessment in the beginning, along with education and providing tools, (3) therapy with minimal contact, which includes communication to a lesser degree, for example via email, and (4) regular therapist-administered therapy. Postel, Haan, Huurne, Becker, and Jong (2010) highlighted the lack of active involvement of the therapist in online delivered interventions and they are now trying to fill in the gap in this field.

Roesler (2017, p. 376) described virtual interaction of the therapist and the client as "a complex mixture of proximity and distance, of presence and absence, of reality and fantasy". In spite of the various changes in therapeutic setting and type of interaction, online treatments seem to have found their place in digital surrounding. Furthermore, many studies have already shown promising results in establishing therapeutic alliance, empathy, and overall achieving positive outcomes in therapy via web-based programmes (Caragea, 2017; Fast, Chen, & Bernstein, 2016; Gibson et al., 2016; Gibson, Malandrakis, Romero, Atkins, & Narayanan, 2015). Online

[1] eHealth and telehealth are used interchangeably throughout the literature

treatments that have made notable progress fall under the category of addictions, especially in domain of smoking (Etter & Geneva, 2005; Ubhi et al., 2016), eating disorders (Low et al., 2006), and alcoholism (Copeland, Ph, & Martin, 2004).

The most encouraging outcomes were derived from interventions that are focused on problem drinkers, in terms of decrease in alcohol consumption during and after the programme (Postel, Haan, & Jong, 2010). Cunningham and Breslin (2004) showed that most problem drinkers never seek treatment, which is why the widespread use of internet could be a fruitful field for developing interventions for this population (Postel, Haan, & Jong, 2010). Furthermore, such Internet health interventions seem to reach populations that are usually harder to involve in face to face care (Humphreys & Klaw, 2001; Postel, Haan, & Jong, 2010), with emphasis on "women, higher-educated people, employees and elderly" (Postel, De Jong, & De Haan, 2005, p. 2393). Motivational Interviewing (MI) with combination of Cognitive Behavioural Therapy (CBT) showed to be efficacious in treatment of substance disorders online (Lundahl & Burke, 2009). Since these treatments mostly rely on client-clinician speech exchange and encouraging client's own verbalization of motivation (Miller & Rose, 2010), newly developed text analysis techniques have been utilized to bestow more insights in how to improve their fidelity and delivery. MI is a counselling style that follows protocols in order to attend to changes in language with purpose of reaching set goals (as cited in Lord, Baer, & Atkins, 2015). These findings indicate that matching styles of communication could lead to positive changes within the therapeutic process. Moreover, newly developed text mining techniques allowed the researchers to identify different categories, thinking processes given between the lines, and follow how exchanged conveyance evolves over the time. The question that may arise is if insights gained using text mining techniques could lead to more comprehension about how

therapists and clients communicate within the intervention, or how their communication styles develop during the treatment.

**Text mining techniques**

Text mining can be defined as a methodological framework that encompasses numerous techniques, which could be automatized and utilized to analyse great amounts of text (Bellot, Bonnefoy, Bouvier, & Duvert, 2013). Since the linguistic analysis was introduced to the field of psychotherapy, research has been focused on either verbal materials classified into theoretically predefined categories based on the theme or semantic content, or on pattern analysis categories based on inherently given structures of the natural language (such as articles, pronouns, abstract words, negatives, function words, etc.) (Halfon, 2017).

For the last decade, text mining has been applied in the domain of sentiment analysis (Mohammad, 2015), content analysis (Salvatore et al., 2012), gaining deeper understanding of process in web-psychotherapy (Murphy, 2015), identifying patterns of effective communication (Howes & Purver, 2012) and detecting empathy in both clients and therapists (Gibson et al., 2016, 2015; Xiao, Imel, Georgiou, Atkins, & Narayanan, 2017). For example, McCarthy, Caputi, and Grenyer (2017) used linguistic processing technologies to explore emotion and cognition, and its relevance to significant change events in different types of psychotherapy, determining that both emotional arousal and cognitive reflection are necessary for optimal emotional processing. Howes and Purver (2012) found that identification of topics, sentiments, and emotion could be used to predict symptom severity and patient's progress with reasonable degrees of accuracy among those who participated in online text-based therapy for depression. Using neural network models, Gibson et al., (2016) managed to predict counsellor's level of empathy ratings from transcripts.

The various ways of employing text mining techniques demonstrated that there is a vast field of application for such tools that can be specifically tailored for psychotherapeutic processes. One of the techniques with the most widespread use is Linguistic Inquiry and Word Count (LIWC), developed by Pennebaker, Boyd, Jordan, and Blackburn (2015). LIWC is a text mining technique that allows exploration of emotional, cognitive and structural components that could be identified in individuals' verbal and written speech (Pennebaker et al., 2015). Due to the variety of categories that are encompassed with LIWC dictionaries, its application has been empirically tested throughout different domains. LIWC has tapped into emotionality, thinking styles, affective states, individual differences, intentions, and motivations (Tausczik & Pennebaker, 2010)

Contemporary research undoubtedly managed to set the cornerstone for further development of linguistic analysis for psychotherapeutic purposes. However, how to utilize this modern approach into therapeutic setting in the most suitable way still requires closer examination.

**Emotion-abstract language**

The way in which emotional content and thinking processes are conveyed using language has also become measurable and available for more thorough analysis. Sassaroli, Brambilla, Cislaghi, Colombo, and Centorame (2014) showed the importance of reflective thinking, cognitive, emotional and interpersonal elements in therapeutic process. Conducting research on Cognitive Behavioural Interventions, they have found that abstract thinking, reflective cognition, and emotional engagement are connected with integrated patterns of therapeutic changes that occur throughout the process. This is in concordance with Therapeutic Cycle Model, established by Mergenthaler (1996, 2008, 2017) in which crucial events of therapy can be pinpointed by

identifying emotion-abstract patterns within the verbalization of the clients. Emotion-abstract patterns can be manifested through relaxing, reflecting, experiencing and connecting, all of them implying different processes that can be recognized from predominating type of language that is used within the session (McCarthy, 2015; Mergenthaler, 1996). Independently from the preferred psychotherapeutic approach, recognizing such patterns has shown that their sequential emergence correlates with positive outcomes in therapy and overall rapport established between the client and the therapist.

In order to empirically grasp processes that manifest evolution of the therapeutic change, Bucci and Maskit (2007, 2015) developed the concept of Referential Process. Bucci and Maskit (2007, 2015) refer to Referential Process as a way of conveying sensory, visceral, and emotional experiences by connecting words to symbolic information. Bucci and Murphy (2015) reflect on Arousal, Symbolizing, and Reorganizing Phases as core processes that lead to operationalization of the Referential Process as Referential Activity (RA). The internal psychological processes, encompassing desires, memories, perception and needs became quantifiable by measuring RA (Bucci, 2013). Changes of the three core phases of RA come in linguistic patterns exchanged in talk therapies, and represent how experiences are expressed and resolved in therapy.

Their study of identifying the RA values in communicational exchange contributed to developing a computer programme named The Discourse Attribute Analysis Programme (DAAP), which tracks the evolution of these processes during the session. DAAP is reported throughout the literature as an useful tool for identifying crucial moments in therapy that lead to synchrony with the therapist and positive outcomes (Bucci & Maskit, 2007; Bucci & Murphy, 2015; Murphy, 2015). Until now, all studies that used DAAP to measure RA were conducted on transcriptions of therapeutic sessions.

**Research question**

Detecting linguistic markers in aggregated text that is collected from online interventions, especially in its efficacy in treating addictions with emphasis on drinking behaviour, show great potential in improving web-based delivered services.

Bearing in mind that language attuning, rapport, and empathy have been perceived as key predictors of positive outcomes in therapy (Gibson et al., 2015; Lord et al., 2015; Lord, Sheng, Imel, Baer, & Atkins, 2016), the field of research is now addressing the matters of how to establish such profound relationship via digital communication.

Since many web-based interventions that use MI and CBT in treating addictions have already found their places in the virtual world, and that newly developed text mining techniques show that reading between the lines and understanding emotion and function words are psychological cues of emotional, cognitive, motivation, and intention states (Tausczik & Pennebaker, 2010), the question that may arise is if the deeper insight into the language used in web-based interventions could provide a new outlook that could improve the online therapeutic process.

Understanding the underlying messages has been a challenge for researchers for quite some time now, especially when it comes to shades of meaning of emotions (Barrett, 2004; Fast et al., 2016). This is especially significant in the virtual communication, due to familiar effect of less inhibited display of behaviours and affects in digital surroundings (McKenna & Green, 2002). Furthermore, all the emerging changes in life and experiences are connected through language to sensory and emotional expression, which are required for functioning well (Bucci & Murphy, 2015).

Due to the fact that emotion and abstract language correspond with successful outcomes in therapy, and that newly developed mining techniques allow identification of such language, the research goal of this paper is to explore the following questions:

- Which categories relevant to emotion-abstract language in web-based interventions for problem drinkers can be identified using LIWC text mining technique?
- Does DAAP text mining technique show Referential Activity within both clients and counsellors in web-based interventions?
- Do LIWC and DAAP show overlapping results in terms of category of the words and evolution of psychotherapeutic process?
- Do results obtained using LIWC and DAAP give sufficient grounds for further research of emotion-abstract language in virtual setting?

## Methods

The goal of this exploratory study is to identify emotion-abstract language using LIWC and DAAP text mining techniques. For the analysis, existing LIWC and DAAP programmes have been integrated into a new software tool.

In order to achieve so, the data used for research purposes was provided from an international sample of "Look at your drinking" (AlcoholDeBaas) online intervention that has been established in the Netherlands and previously shown promising results (Postel, Haan, Huurne, et al., 2010; Postel, Haan, & Jong, 2010). The data consists of exchanged emails from both counsellors and clients during the web-based treatment, and drinking diary logs which were filled in by clients. Since the study encompasses a sample of nine participants, employing mixed methods was possible. This included gathering results from LIWC and DAAP text mining

technique, and in-depth qualitative analysis of individual cases. Qualitative analysis of the data enables the researchers and readers to gain deeper and clearer understanding of the reasoning behind retrieved quantitative data and the timeline of each case. Such mixed approach provides context, interpretation, and better understanding value of the quantitative results obtained with LIWC and DAAP.

The first part of this section provides a reader with the description of the intervention. This overview was included in the study in order to introduce readers to the structure of the programme, and to encompass all topics that were covered in the intervention. In this web-based intervention, counsellors often sent questionnaires or tailored emails, in which context of the message is adjusted to either client's needs, or according to the part of the programme. Since the data consists of the emails exchanged by counsellors and clients, list of the topics and structure of the intervention gives a necessary context for data analysis and helps interpretation of the findings.

**Look at your drinking**

"Look at your drinking" is an international adaptation of AlcoholDeBaas programme, which was established in the Netherlands in March 2015 (Postel, Haan, & Jong, 2010). According to the study of Postel, De Jong, and De Haan (2005), the treatment was intended for those who are concerned about their drinking patterns. The patients in the treatment were referred by their own general practitioner. Participants were eligible if they were above sixteen years old. The program was delivered completely online, via the Internet website and asynchronous email exchange between the therapist and the user. This means that pace of email exchange was based on how fast client responds, and on client's progress in terms of parts of the treatment. Counsellors usually respond within three working days, while sometimes include

useful articles in the meantime. The fact that therapist was actively involved during the whole intervention distinguishes this online interventions from others, allowing users to have both support and guidance throughout the process. The intervention is based on the "cognitive behaviour therapy (Hester, Miller, & Goldman, 1996; Irvin,1999), motivational interviewing (Britt, Hudson, & Blampied, 2004; Miller & Rollnick, 2002), and Prochaska and DiClemente's (1983) Stages of Change Model, all empirically supported methods for substance-use disorders in regular face-to-face addiction treatment" (as cited in Postel, Haan, & Jong, 2010). Such integrative approach is common in online interventions, due to their structure and coding schemas, which makes it easy to deliver, measure, and assess. Nonetheless, Motivational Interviewing (MI) have been successfully assessed for identifying empathy in online interventions via language synchrony (Lord et al., 2015), predicting empathy and behaviour (Gibson et al., 2015; Mihalcea et al., 2014)

The intervention itself is divided in two main parts. The first part entails exploration of drinking habits throughout two assessments and four assignments, including:

- Exploring advantages and disadvantages
- Understanding of drinking patterns
- Filling in the drinking diary
- Identifying the situations in which drinking is mostly likely to occur

In the second part, after the consultation and recommendation of a multidisciplinary team whether the participant should continue with the treatment, the client is involved in the following:

- Setting a drinking goal
- Formulating helpful and non-helpful thought,

- Considering helpful behaviours for moments of craving,
- Identifying the moment of the decision to drink alcohol,
- Formulating an action plan for maintaining the new drinking behaviour and for relapse prevention. (Postel, 2011)

The intervention lasts on average sixteen weeks and allows setting your own pace. The care provider contacted patients once per week and he or she was obliged to answer within three working days. The content of the emails is standardized and informative in its nature (Roskam, 2013). The therapist still has to personalize the messages and address emerging matters in mutual correspondence. After the treatment, clients are given an opportunity to opt for Finger-on-pulse services and follow-ups. This includes aftercare and keeping contact with the assigned counsellor. This part also includes receiving questionnaires regarding the satisfaction of the treatment, sent six weeks and half a year after the treatment was completed.

**Linguistic Inquiry and Word Count (LIWC)**

LIWC is a text mining technique that allows exploration of emotional, cognitive and structural components that could be identified in individuals' verbal and written speech (Pennebaker et al., 2015). This method is often used in combination with Motivational Interviewing and CBT studies (Mihalcea et al., 2014; Tanana et al., 2017). LIWC "uses counts of words within user-defined segments, such as turns of speech, that match words in dictionaries defined by particular grammatical or psychological categories" (Murphy, 2015, p. 135). The strategy behind the word count is based on the assumption that psychological information could be conveyed throughout the used words, independently from the semantic context and literal meaning (Pennebaker, Mehl, & Niederhoffer, 2003). The potential of grasping emotional and cognitive processes from the written or verbal material with LIWC software has been recognized

in various studies: oral history (Truong, Westerhof, Lamers, Jong, & Sools, 2013; Westerhof, Lamers, & Jong, 2014), empathy and language synchrony (Lord et al., 2015), online support groups (Haug et al., 2008), and more.

LIWC computes the number of occurrences of words that are divided in 75 categories, which allow overlapping. These categories include standard linguistic dimensions (e.g. pronouns, articles, auxiliary verbs), psychological constructs (cognitive, biological processes, affect, drives), personal categories (work, home, free time), and punctuation categories (periods, commas, quotation marks) (Pennebaker et al., 2015). Since the focus of this study is on the emotion-abstract language, the categories have been carefully assessed in order to identify those that were relevant for the research purpose. The umbrella category of Psychological Processes was found as applicable. This category consists of: cognitive processes, social processes, affective processes, perceptual processes, and personal concerns. Personal concerns do not comprise out of one umbrella category, and it is divided in subcategories including work, leisure, home, money, religion, and death.

The primary focus is on categories of cognitive processes and affect, due to the nature of the categorized words in the LIWC manual. These categories encompass words that are relevant to thinking, insights, inhibitions, positive and negative emotions, and more specific emotions such as sadness and anger. Other categories were included due to the expectation that the topics they cover would be frequently mentioned or addressed in the intervention, and that they could be connected with thinking processes and feelings.

Examples of words that are associated with the selected categories are presented in Table 1.

Table 1

*Examples of words in chosen LIWC categories*

| Category | Word examples |
|---|---|
| Social processes | Daughter, husband, buddy, friend, baby, adult, neighbour |
| Affective processes | Happy, cried, abandon, love, nice, sweet, worried, annoyed, crying, grief, sad |
| Cognitive processes | Cause, know, ought, effect, should, would, could, maybe, guess, always, never, stop, include, without |
| Perceptual processes | Observing, heard, feeling, listen, hearing, feel, view |
| Biological processes | Eat, blood, pain, spit, hands, clinic, flu, love, horny, dish, eat |
| Personal concerns | Job, earn, win, cook, chat, apartment, family, cash, owe, church, bury, coffin |

*Note.* Reprinted from "The Development and Psychometric Properties of LIWC2007", by Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K., 2015, p. 5. Austin, TX: University of Texas at Austin.

**Discourse Analysis Attribute Program (DAAP)**

Discourse Attributes Analysis Program (DAAP) analyses great corpora of text in order to track occurrence of narration and imagery with the aim of finding linguistic predictors of descriptions and reflections on their meaning (DAAP; Bucci and Maskit 2005; Maskit, Kingsley, and Welsh 2006; Maskit and Bucci 2006). In addition to the usual linguistic analysis functions, such as computing mean usage of words from specified dictionaries, this program uses a mathematically smooth weighting function to compute an average of the dictionary matches at given times. DAAP has been tested with different materials and showed limited, but promising results, especially within the field of psychotherapy and evolution of psychotherapy treatment. In the field of psychotherapy, DAAP has demonstrated the ability to track the change of the language being used throughout the therapeutic sessions (Murphy, 2015).

Bucci and Murphy (2015) argued that changes that occur during the sessions are manifested in verbal exchange and identifiable with Referential Activity (RA). The core processes that underpin RA are Arousal, Symbolizing Phase, and Reorganizing Phase. Arousal Phase comprises of bodily sensations, or plans for motor action that have not yet passed the threshold of verbalization, which causes greater disfluency in speech (Bucci & Murphy, 2015). Symbolizing Phase includes preliminary par when the material reaches symbolic form, and the latter part in which the symbols could be reflected in words, making them more specific and easier to comprehend by reader or listener (Bucci & Murphy, 2015). The final phase is more focused on cognitive processing, logical operations, and reasoning, allowing the speaker to reflect and reorganize their thoughts and emotions. Since RA encompasses different processes that employ abstract language and emotional expressions, which are measurable via linguistic markers within the DAAP programme, identifying moments of higher activation of these phases imply use of emotion-abstract language. DAAP programme's dictionaries are based on the assumption that underlying psychological processes are represented in the words provided, and that variations of language within the session manifest these processes (Murphy, 2015).

The available dictionary for this study was WRAD, weighted dictionary that reflects to which extent Referential Activity (RA) was represented in the study. WRAD dictionary contains 696 items in total and weights them from -1 to 1, in order to give the overall RA activity output (Bucci & Maskit, 2006). "For example, an item with weight −1 is used much more often in text segments having RA scores in the range of 0 to 2.75, an item with weight +1 is used much more often in text segments having RA scores in the range of 7.25 to 10" (The Discourse Attributes Analysis Program DAAP , 2019). RA indicates how emotional, bodily, and imagery experiences are reflected in language, and if they evoke corresponding response within the listener or a reader

(Bucci, Maskit, & Murphy, 2016). Figure 1 contains an example in which highly weighted RA words are given in dark red, and low RA in dark blue, while other colours indicate neutral values.

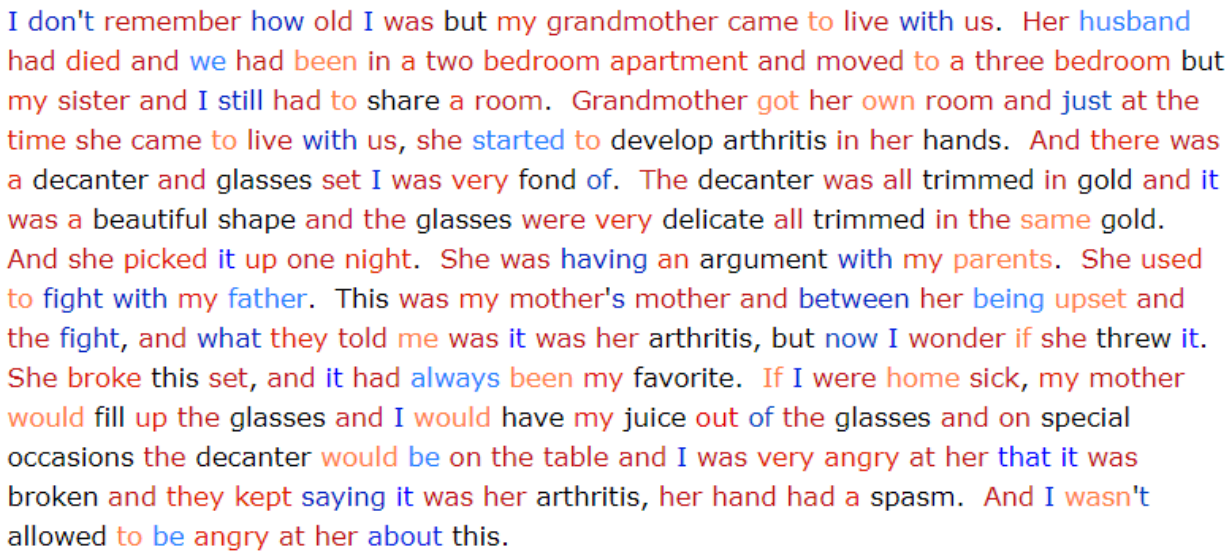



*Figure 1.* Example of RA activity as seen with WRAD dictionary

*Note.* Reprinted from The Weighted Referential Activity Dictionary (WRAD). (2019, March 6). Retrieved from The Referential Process: http://www.thereferentialprocess.org/dictionary-measures-and-computer-programs/weighted-referential-activity-dictionary-wrad

**Ethical approval**

Cases used for this study have been retrieved from Tactus addiction care and its program Look at Your Drinking, which is an international adaptation of the Dutch web-based intervention AlcoholDeBass. In order to respect privacy of the available clients, the data has been completely anonymised. During this process, all personal information that could identify a person has been removed. The data was available only to the researcher and the programmers who analysed it. The data was carried on an encrypted and password secured USB stick. The short summaries of each cases that are provided in this paper do not compromise the privacy of the participants. The descriptions are included in order to provide the readers with better understanding of the context in which language was utilized.

Since the data is collected from the Tactus Addiction Care Centre, the participants gave informed consent at the beginning of the treatment. The ethical grounds were set and approved by the scientific committee of Tactus Addiction Care, which carries responsibility for ethical issues.

**Description of the data**

The data consists of the written exchange between the therapist and client, which is collected from emails and drinking logs. The emails that were sent by the counsellor are either standardized and premade, or tailored and adjusted accordingly towards the users, depending on the situation and the content of the correspondence. Topics that were provided in the description of the intervention "Look at your drinking" are encompassed in premade pool of messages. Diary logs were filled in by clients and they include information about the amount of the urge to drink, situation when the urge occurred, and emotions and thoughts that client experienced at that moment. The reasoning behind including the information from the drinking logs is the observed discrepancies between the reported emotions and intentions within both types of data. In order to provide an overview to the reader of the cases, all nine clients from the sample will be summarized and described using the following characteristics:

- The alcohol use per week at the beginning of the treatments
- The goal that the client has set
- Other reported personal problems
- Alcohol (or other addiction, such as smoking) history
- General responsiveness to messages
- Treatment satisfaction and goal reached
- Observed discrepancies between the intake and reported messages

In the description of the cases, alcohol consumption is expressed in consumed alcohol units. This measurement quantifies the alcoholic content in a beverage, as an approximate guideline for total alcohol consumption (Unit of alcohol, n.d.). The measurements are modified in order to adjust the fixed amount of pure alcohol with the concentration of the beverage, meaning that units differ depending on the type of alcohol (Standard drink, n.d.). For example, a medium glass of wine contains around two units of alcohol, whereas a typical glass of beer contains almost exactly one unit (Unit of alcohol, n.d.).

**Preprocessing of the data**. In text mining methods, certain steps need to be taken in order to make the data ready for analysis. In order to do so, all texts from the emails were anonymised and normalised. Anonimisation was conducted in order to remove any personal information from the data (names, birth dates, countries of residence, companies in which clients work) that could compromise the identity of the clients (Tjong et al., 2019). Normalisation process removes all capitals, italics, and bold text forms, making the data readable for the employed programmes. The texts for analysis were derived from emails, meaning that questionnaires received after treatment are excluded. Additionally, logs from diaries were included in analysis in order to grasp more information of emotion and cognition of the participants.

Before both DAAP and LIWC are used, there were additional alternations made to the data. The reasoning behind the action lies in the attempt to make data retrieved from emails more valuable for the study. The following changes were made:

1. Most of the clients used the “reply” option when responding, which meant both answers from the client and the counsellor were included in the client’s email. Since this had data skewing as a consequence, the forwarded emails from the counsellor were not included in further analysis.

2. The sample of the study included participants from the United Kingdom, which is why word “fancy” was included in the analysis. This word is frequently used to denote liking, preferring, or desiring something.
3. Two out of nine participants responded to counsellors’ emails with forwarding their questionnaires back, while marking their answers in italics, capitals, or bold text. Since the bold and italics markings were unavailable due to preprocessing, only texts in capital were used for analysis. The text in capitals has been copied in the clients’ emails, when possible.

## Results

This section includes results and findings of this study. In the first part, a brief description of individual cases is provided. The second part includes general information about the whole sample and the dominant categories that were identified using LIWC text mining technique. The third part contains of presentation of one case out of the whole sample, displaying insights retrieved using LIWC and DAAP techniques. Last part provides an overview and summary of the whole result section.

### Case description

In this part, cases will be presented and summarized according to the previously set characteristics. The aim of this section is to provide the readers with an overview of each individual case and important categories that were analysed using the text mining techniques. All case descriptions include information that do not compromise privacy of participants. Personal information was omitted with anonimisation during preprocessing.

**Case 1.** The client is a 38 year old female with a 10 year history of alcohol consumption. At the beginning the treatment, client was consuming 62 units per week. The goal that client set

was to reduce alcohol intake in order to prevent health problems and progress in her studies. The main motivation for drinking includes its rewarding and relaxing effect. In terms of other health related issues, client reports having problems with weight (already taking B supplements). Previous drinking habits have caused her loss of memory, missing lectures, self-sabotaging patterns (in terms of diet), headaches, sleeping issues. An important factor for this client is also her husband, with whom she drinks on many occasions. The client notices problems quitting even when others around her are not drinking.

**Case 2.** The client is a 49 years old female. At the beginning of the treatment, client consumed 55 units of alcohol per week and she set a goal of reducing the consumption. The client reported having history of depression, suicidal ideations, but she states does not have them at the moment of the treatment. There are evidence of isolation and depressive thoughts. She is willing to begin exercising and improve her routines. The client has previous dry periods for two years after the therapy. The fact that she self-detoxed before twice were the reason for therapeutic interventions. The main motivators for drinking encompass relieving anxiety. The most dominant feelings in the emails and diaries include religious outlooks, feelings of shame, disappointment, remorse, and frustration. The client is aware of her negative cycles and she was willing to employ adopted behavioural and cognitive strategies. The client perceives alcohol as a way of facing problems, since she lacks the tools to cope with them by herself.

**Case 3.** The client is a 49 year old male, living with his wife and 2 children. The level of consumption at the beginning of the treatment included between 20 (self-reported) and 50 (self-reported) units per week. The client set the goal of reducing consumption, but it emphasizes the social and practical factors that justify the drinking behaviour. Physical symptoms that were noticed are memory problems, tiredness, sexual problems, depression, and mood swings.

Important factor during the treatment is client's wife, who is described as a "regular drinker", and she is often mentioned in diary logs as an enabling factor. During the treatment, mood swings were dominant. Due to the nature of the client's job, alcohol could be understood as an important contributor of settling the business deals. Mood swings that were frequently reported in the emails are especially emphasized in the log data, in which alcohol is perceived as a reward.

**Case 4.** Client is a 38 year old male, living with his partner and his son. Consumption at the starting point included 30.8 units of alcohol per week. Drinking for this client is an omnipresent problem since his early years and even in the family setting. The main reason for continuing with this lies in relaxing and alleviating affect that alcohol has on this client. For this client, alcohol is also a social factor, and a habitual response in a friendly/family setting. The client reported having headaches that might be a consequence of large amount of beverages, but he decided not to share this information with his GP. The initial goal to reduce the consumption was met. The most often reported emotions by the client include embarrassment from the partner and rewarding characteristics of the consumed alcohol. Additionally, the client reported experiencing cravings, fatigue, and mood swings, caused by reduced intake. The client withheld information from his family and friends, due to above mentioned reasons.

**Case 5.** Client is a 27 year married female, who got pregnant during the treatment. The client has medical history of potential eating disorders (reports binge eating, dieting, excessive exercises). Physical symptoms include fatigue, headaches, and sleep deprivation. At the beginning of the treatment, client consumed 32 units of alcohol per week. Previous operation indicates drinking and dieting problems. Alcohol is perceived as a rewarding, socializing, and self-empowering factor. Due to the working environment, the client feels as if it is mandatory

that she does not deny an offered drink. On emotional level, client refers to disappointment and frustration, which are related to her self-esteem problems that are also reflected in her problems with body-image and previous binge eating episodes. For this client, establishing control over her symptoms are of great importance. Most of the diary log information are consistent with the reported consumptions, but there are justifications, such as “I deserve this drink”. Due to pregnancy, the client decided to alter her goal from reduction to complete abstinence.

**Case 6.** The client is a 59 year old female and identifies as a “loner”. At the registration point, the client consumed 41.3 units per week. She lives by herself and she is educated in Korean Master Yoga meditation, which helped her in recovering from head injury caused by an accident. Regular medication did not help, due to the fact that the client is allergic to many forms of medication. In her treatment, the client refers to her spiritual path that keeps her going and reduce her drinking. In the emails, client emphasize that she understands the long process that is ahead of her, and she tries to find additional support in her spiritual beliefs and her own capacities. The main reasons for drinking are related to numbing the feelings and reducing emotional pain. She has a history of starting consuming alcohol early, but with long dry periods. The client frequently consumes alcohol in the evenings and on the weekends. The client is introspective in terms of her own contributions to drinking patterns and often connects her determination with the spiritual path that she wants to follow. Most of the reported drinking occasions are related to spontaneity or to justify her reasons for acting on a habit.

**Case 7.** The client is a 37 year old divorced female, who lives with her 2 sons. The client is currently in a relationship with who she plans to get married. At the beginning of the treatment, the client consumed between 50.8 units per week. The alcohol is perceived as a reward and emotional support during the occasional bad moments or fights with her future

husband. The client reports having low self-esteem from her divorce and inability to cope with extra work load and family issues. During the treatment, the client did not share her drinking pattern or extent of her problems with her partner. Additionally, the partner gave her an ultimatum to stop consuming alcohol or they will not get married. During the treatment, the client was mostly focused on getting advice and not acting on it, which has reflected in her unchanged levels of consumption throughout the treatment. She often reports of "not seeing the point" of trying to alter her behaviour.

**Case 8.** The client is a 48 year old male, living alone and conflictual relationship with his sister. His mother is a nursing home, and he often interacts only with one friend. The initial consumption included 105.6 units per week. The client reported having a breakdown, due to which he consumes medication, in addition with blood pressure and cholesterol ones. The most dominant emotions in the emails are anxiety, frustration, and overall dissatisfaction with his relationships, work, and life. The client shows indications of suicidal ideation and is often apologetic towards the counsellor. He perceived alcohol as a social factor, highlighting that "there is no life or parties without alcohol".

**Case 9.** The client is a 42 year old female who lives alone and reports having a relationship with a boyfriend. The initial consumption included 63 units per week, which are in discrepancy with 98 units that are recorded in her log book. The main reason for consumption lies in socializing. When the client is alone, she is capable of dry periods, which she included in her treatment. She often cured hangovers with alcohol and reports consummation as normalized in the family. Additionally, the client would gamble once per week, but she did not reveal further information. The client suffers from tiredness, fatigue, joint pains, nausea, exceeding sweating, and memory problems. Her weight is affected by binge eating episodes. Dominant emotions are

depression, despondence, and awareness of the extent of her problem. The client actively employed adopted behavioural and cognitive strategies.

**Sample overview**

Nine participants in total were included in this study. Six participants were female, three male, ranging in age from 27 to 59. The average age of the participants was 43 (SD=33), and the consumption was 54.5 (SD=22.41). In the Table 2, gender, age, and the initial consumption per week are shown.

Table 2

*Gender, age, and consumption of the participants*

| Gender | Age | Consumption at the beginning of the treatments |
|---|---|---|
| Female | 38 | 62 |
| Female | 49 | 55 |
| Male | 49 | 50 |
| Male | 38 | 30.8 |
| Female | 27 | 32 |
| Female | 59 | 41.3 |
| Female | 37 | 50.8 |
| Male | 48 | 105.6 |
| Female | 42 | 63 |
| M | 43 | 54.5 |
| SD | 9.33 | 22.41 |

All participants proceeded to the second part of the treatment, among which only one participant was allowed to do so under specific conditions. None of the clients signed for Finger-on-pulse option. Due to the fact that only one participant completed the intervention, information regarding consumption at the end of the treatment is not available for all cases. In most of the cases, the treatment ended with inactivity of the client and/or not responding to the last couple of emails.

Email correspondence between the participants and counsellor is unequally distributed. On average, clients sent 15.4 emails (SD=4.85), whereas counsellors sent 31.5 (SD=6.04). In sum, the whole pool of messages include 139 emails sent by participants, and 284 emails sent by counsellors. Counsellors always exceeded the number of client's sent messages. Table 3 contains the overview of emails and diary entries for all 9 cases.

Table 3

*Exchanged emails and diary logs*

| Data set | Client | Counsellor | Diary |
|---|---|---|---|
| 1 | 7 | 21 | 15 |
| 2 | 13 | 31 | 26 |
| 3 | 17 | 33 | 141 |
| 4 | 15 | 26 | 33 |
| 5 | 13 | 35 | 25 |
| 6 | 23 | 43 | 41 |
| 7 | 12 | 33 | 27 |
| 8 | 20 | 30 | 77 |
| 9 | 19 | 32 | 16 |
| Sum | 139 | 284 | 401 |
| M | 15.44 | 31.56 | 44.56 |
| SD | 4.85 | 6.04 | 40.65 |

In terms of diary data, the differences among cases vary from 15 to 141 inputs, making the sum of all data available 401. Clients did not fill in their diaries consistently, often leaving gaps or completing the inquiries retroactively. This prevents estimating decreases or increases in alcohol use.

Table 4 encompasses tokens (numbers of words) that were exchanged by clients and counsellors during the whole treatment for all nine cases. On average, clients used 4414.4 (SD=2605.34) words, whereas counsellors used 17076.4 (SD=3843.17) words.

Table 4

*Clients' and counsellors' tokens*

| Data set | Clients' Token Average | Counsellors' Token Average |
|---|---|---|
| 1 | 1996 | 12627 |
| 2 | 3915 | 19886 |
| 3 | 4784 | 19959 |
| 4 | 2795 | 13963 |
| 5 | 1842 | 19750 |
| 6 | 3946 | 19141 |
| 7 | 4885 | 18791 |
| 8 | 4983 | 9906 |
| 9 | 10584 | 19665 |
| Sum | 39730 | 153688 |
| M | 4414.4 | 17076.4 |
| SD | 2605.34 | 3843.17 |

**Results obtained using LIWC.** The preselected LIWC categories were examined among both client and counsellor in order to compare and contrast how these topics emerged between these two correspondents. The categories are calculated as relative to the total tokens used by either a client, or a counsellor. This implies that results are not comparable on the absolute scale, but they provide the reader with the overview of how the categories were distributed in the overall language use of both senders. All tables per correspondence are available in the appendices, whereas in this section general overview will be provided. Most dominant categories throughout the sample were cognitive, affective, and social processes.

Table 5 shows the percentage of the prevalence of the category for most dominant topics.

Table 5

*Clients' and counsellors' Cognitive, affective, and social processes*

| Data set | Clients | | | Counsellors | | |
|---|---|---|---|---|---|---|
| Processes | Cognitive (%) | Affect (%) | Social (%) | Cognitive (%) | Affect (%) | Social (%) |
| 1 | 7.57 | 5.06 | 4.26 | 10.53 | 6.63 | 10.02 |
| 2 | 11.57 | 6.85 | 5.29 | 9.94 | 6.79 | 10 |
| 3 | 8.65 | 5.48 | 4.31 | 9.58 | 6.73 | 9.99 |
| 4 | 8.62 | 4.47 | 4.54 | 9.70 | 6.35 | 9.98 |
| 5 | 5.86 | 4.02 | 2.44 | 10.70 | 6.54 | 10.25 |
| 6 | 8.34 | 4.69 | 3.52 | 9.94 | 6.50 | 10.29 |
| 7 | 9.79 | 5.71 | 6.45 | 9.71 | 6.57 | 10.22 |
| 8 | 9.75 | 5.26 | 7.79 | 8.27 | 5.69 | 11.34 |
| 9 | 9.93 | 6.09 | 7.68 | 10.23 | 6.34 | 10.14 |
| M | 8.90 | 5.29 | 5.14 | 9.84 | 6.46 | 10.25 |
| SD | 1.63 | 0.86 | 1.84 | 0.70 | 0.33 | 0.43 |

The high prevalence of cognitive and affective processes across the sample indicates that emotion and cognition related words were utilized within both clients' and counsellor's emails. On average, cognitive processes were used 8.90% (SD=1.63) of the time by the clients, meaning that in their emails they used words related to insights, causation, certainty, inhibition, causation, and tentativeness. The counsellors used cognitive process words 9.84% (SD=0.70) of the time. In most cases, with the exemption for the case one and five, cognitive tokens among both clients and counsellors were within similar values. In terms of the affective processes, clients referred to positive and negative emotions, anxiety, anger, and sadness within 5.29% (SD=1.84) of the emails, whilst counsellors used them in 6.46% (SD=0.33) of the time. Counsellors used affective related words more often consistently throughout the sample, even though the clients referred to this category on less frequent basis. Figures 2 and 3 show the distribution of these categories identified in all emails.

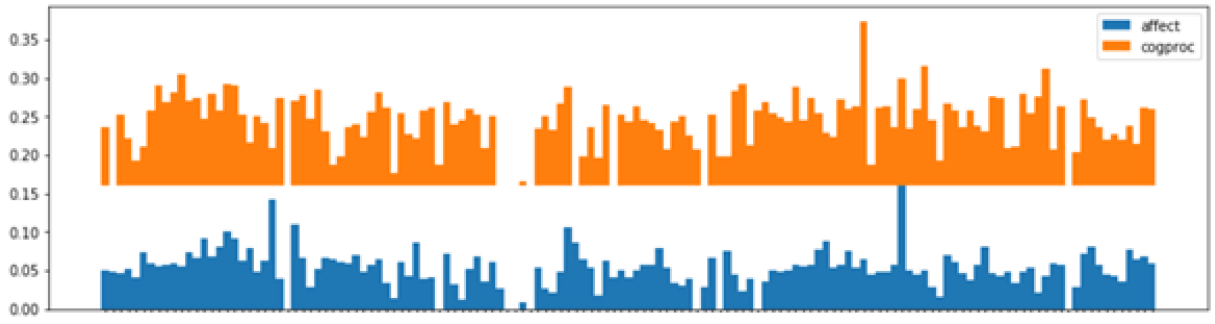



*Figure 2.* Client' cognitive and affective processes

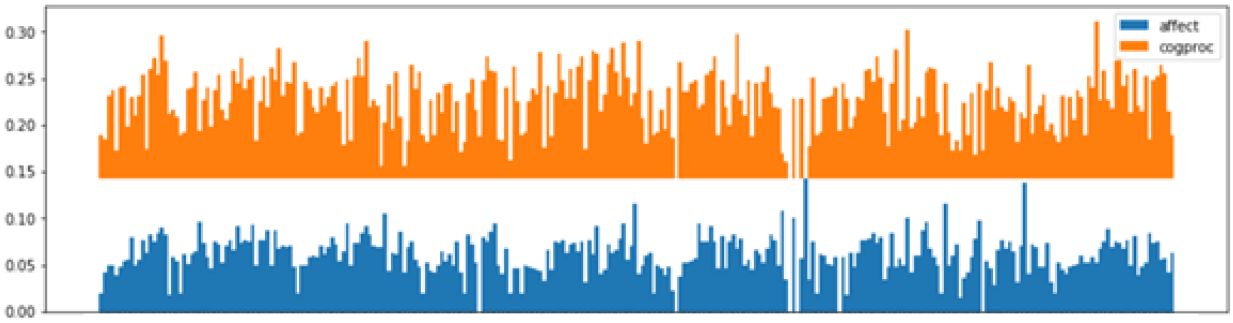



*Figure 3.* Counsellors' cognitive and affective processes

The given graphs demonstrate the distribution of the affective and cognitive processes throughout the time. On the graphs, all categories start at zero value, even though they are portrayed as one above the other. For example, all counsellors' cognitive processes start at value 0, and the maximum value on the graph is around 0.30, meaning that the actual value is 0.15, since the graph for cognitive processes start at 0.15. The counsellors have sent more emails with cognitive and affective content more often than the clients, but the maximum value among clients for both categories are higher. Even though the higher values can be observed among clients (affective processes reach 0.15, and cognitive processes reach 0.20), these results cannot be compared, due to before mentioned scaling of the tokens to relative use. Nonetheless, the distributions of categories per case seem to follow the same logic. In all cases, the same trend has been observed. If specific process was included in clients email, the counsellor addresses that process in the following message, or vice versa (for reference tables and figures are available in

appendices section).

If the results are portrayed on the scatter plot, both senders usually overlap in terms of cognitive and affective processes, as can be seen in Figure 4. The graph presents how both affective processes and cognitive process words that occurred within clients and counsellor cluster. Despite the fact that comparison is not possible in term of exact measures, the scatterplot shows that both senders consistently used cognitive and affective tokens in their emails. The y axis displays words related to cognitive processes, whereas x axis shows presence of affect related words. The blue dots indicate counsellors' emails, and pink ones denote clients' emails.

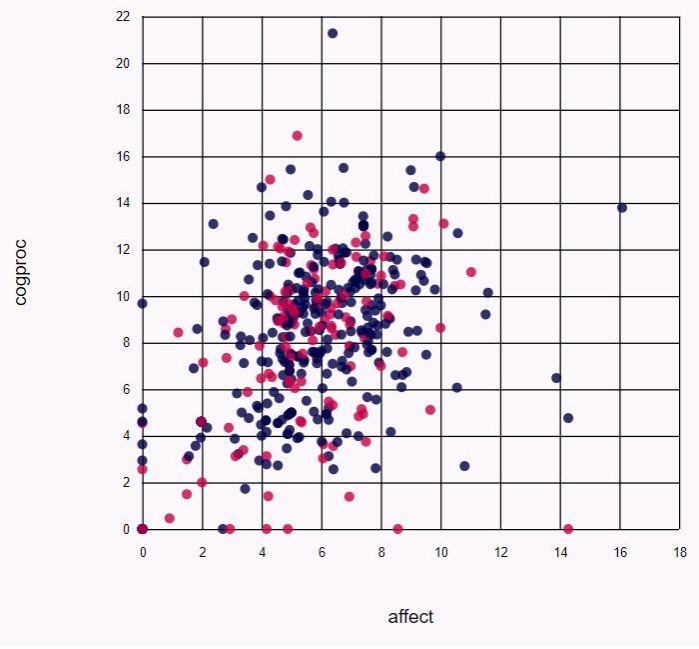



*Figure 4.* Clients' and Counsellors' tokens on scatter plot

Biological and perceptual processes were present in the sample, but not as prevalent as cognitive, affective, and social processes. Subcategories of personal concerns were not always retrieved within the text and none of them have been identified as highly prevalent. Both counsellors and clients referred to work and leisure time more frequently than to other

subcategories. Across the sample, religion, death, and money were the ones that were often missing.

**Results obtained using DAAP.** Referential Activity was present throughout the process in cycles, varying from low to high. In all of the cases, the RA score did not reach a value over 0.025 for clients, and 0.0125 for counsellors.

The higher DAAP scores occurred in the beginning or the middle of the treatment. Detailed DAAP results were interpreted within the selected case.

**Case analysis using LIWC and DAAP**

In this section, one case out of the sample will be presented in order to report results obtained using LIWC and DAAP in order to demonstrate their abilities in detecting linguistic markers of emotion-abstract language in the best way.

For the purpose of presenting results of detecting emotion-abstract language using LIWC and DAAP techniques in web-based intervention “Look at your drinking”, case number two will be analysed. The reason for choosing this particular case lies in the observation of RA values and how the process can be interpreted according to the content of the email. Additionally, all relevant LIWC categories were identified both via analysis and via qualitative interpretation.

Client number two is a 49 year old woman who identifies as a loner. She has history of dry periods, depression, and suicidal ideation. In her emails, she often reflects on how her feelings, thinking patterns, and previous depressive episodes affect her drinking. She perceives alcohol as a way to relieve her anxiety and as a problem-solving method. She recognizes her lack of capacity at the moment to cope with her addiction, but she puts effort in understanding how she could change her behaviour and attitude towards it.

**Results obtained using LIWC.** Client number two sent 13 emails with 3915 tokens, whereas the counsellor sent 31 emails with 19886 tokens. During this correspondence, all categories were identified, with the exception for religious, death-related, and sexual content.

Table 6 shows the categories that were predominantly represented among predefined chosen tokens and categories.

Table 6

*Case 2 – Overview of LIWC categories for case number two*

| Processes | Client (%) | Counsellor (%) |
|---|---|---|
| Affect | 6.85 | 6.79 |
| Biological | 1.63 | 1.64 |
| Cognitive | 11.57 | 9.94 |
| Perceptual | 2.5 | 1.95 |
| Personal concerns | | |
| Work | 0.97 | 1.01 |
| Leisure | 0.92 | 1.38 |
| Home | 0.18 | 0.17 |
| Money | 0.18 | 0.17 |
| Religion | Missing | 0.01 |
| Death | Missing | Missing |
| Social | 5.29 | 10 |

As was reported for all nine cases, cognitive, affective, and social tokens were the most frequently used ones between both counsellor and client. Cognitive processes were used 11.57% of the time, whereas the counsellor responded using this type of content 9.94% of the time. Affective processes, encompassing positive, negative emotions, sadness, anger and anxiety, were mentioned 6.85% of the time by the client. Counsellor reacted to these emotions with same tokens 6.79% of the content in the email. The only exemption for this case lies in social category. Counsellor referred to social tokens 10% of the time, whereas client used them 5.29% of the time.

Biological processes were reported to a lesser extent, 1.64% of the time by the counsellor, and 1.63% of the time by the client. Perceptual category is used 1.60% by the client, and 1.93% of the time by the counsellor, but there were no specific information in the text that could unambiguously be identified with perceptive processes.

Figure 5 portrays how these categories emerged throughout the time for both client and counsellor in case number two.

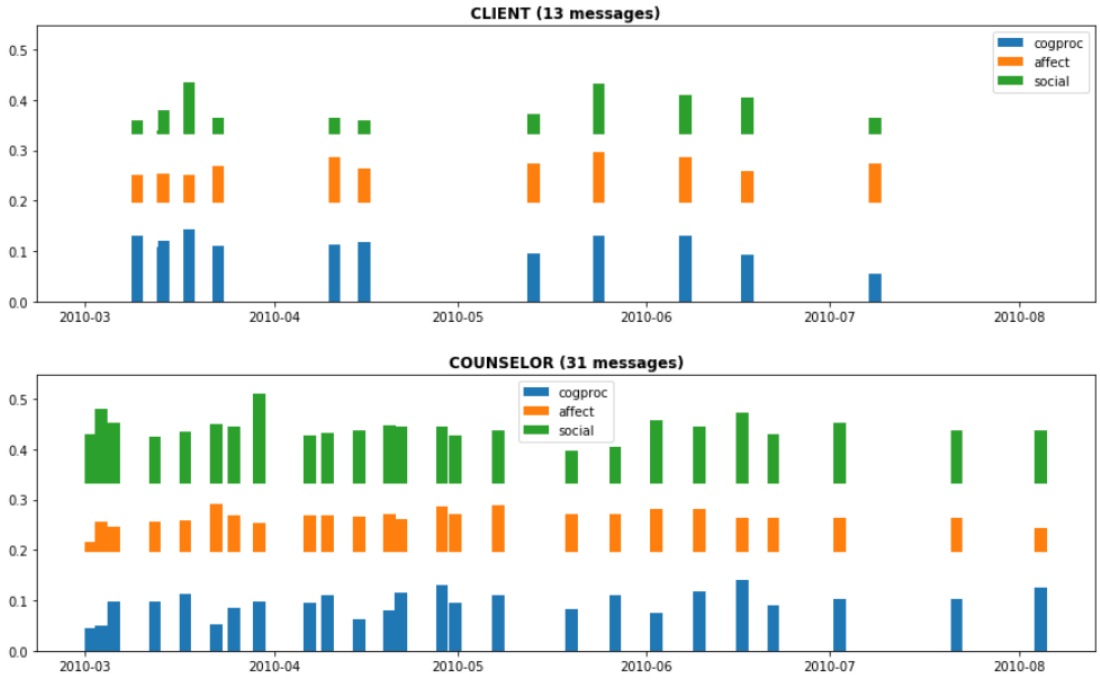



*Figure 5.* Distribution of social, cognitive, and affective processes in case number two

Since case analysis using mixed models allows understanding the context of the given categories, the predetermined tokens were available for identification within the emails. In case 2, social content was around average values, but in the content itself it did not play the usual role it had in other cases. The client identified herself as a loner and she was prone to isolation, which diminished the effect of social factor on drinking habits. Counsellor often offered alternatives or

finding support within close relationships, whereas client usually referred to a meeting with a high school friend she had.

The vastly dominant cognitive category was in concordance with the content of the emails. Client often reflected on how her attitude and the way she perceives drinking affect her current problem. Counsellor consistently addressed emerging thinking patterns as reported by the client. She often reflects on her attitude towards the programme, and how she perceived her goals and progress, endowing the content of the email with problem solving approach.

Affective tokens were reflected in text in identifying feeling of remorse, guilt and fear of relapsing. In terms of affective processes, the most frequently mentioned emotions were guilt and shame (in emails), as well as the alleviating effect of the alcohol on stress and troubling thoughts (in diaries and questionnaires). She often addresses how she feels about the process and reflects on her drinking patterns. In the exchanged email, the counsellor gave priority in addressing client's thoughts and feelings of guilt. Counsellor often showed sympathy for the client, and thanked for the client's engagement in the programme. All the mentioned examples portray how all prevalent categories in this case were presented in the content, but also how overlapping of these categories is inevitable when context is taken into account. Even though the distribution of the categories is not the same, and that the counsellor's emails were more frequent, the values for cognitive and affective processes are similar within the timeframe. For example, client's cognitive and affective language had higher values than counsellors most of the time. In other cases, the counsellors' emails usually surpassed the values in mentioned processes. In case two, even though the counsellor average values were slightly lower than clients, the correspondence between them seemed more consistent and attuned than in other cases.

The client did not report any medical issues that could have fallen into biological category, but she commented on finding additional help within the local clinic, which might have been identified by LIWC within this category. Additionally, counsellor suggested visiting a doctor, but specific reason for this was not mentioned. The similar values within the counsellor for the same category might be due to responding to this comment, or listing common physical (pain, eating habits, and sleep) problems that are a by-product of excessive drinking. Personal concerns were not as prominent as other categories, but both client and counsellor consistently addressed them. For example, client reported having problems at work, and counsellor responding expressing empathy for her job situation.

Table 7 provides readers with overview to which extent subcategories of social, cognitive, and affective processes were the most prevalent ones.

Table 7

*LIWC subcategories for cognitive, affective, and social processes for case number two*

| Data set | Client | | | Counsellor | | |
|---|---|---|---|---|---|---|
| Processes | Cognitive (%) | Affect (%) | Social (%) | Cognitive (%) | Affect (%) | Social (%) |
| Insight | 3.4 | | | 3.47 | | |
| Causation | 1.2 | | | 1.52 | | |
| Discrepancy | 1.66 | | | 1 | | |
| Certainty | 0.97 | | | 0.99 | | |
| Tentative | 2.58 | | | 2.33 | | |
| Positive emotions | | 5.62 | | | 5.45 | |
| Negative emotions | | 1.17 | | | 1.24 | |
| Anxiety | | 0.33 | | | 0.53 | |
| Anger | | 0.1 | | | 0.07 | |
| Sadness | | 0.33 | | | 0.24 | |
| Family | | | 0.18 | | | 0.01 |
| Friends | | | 0.31 | | | 0.11 |

In communication of both client and counsellor, insight and tentative words were used in majority of the written text. The words related to insight subcategory are such as: think, consider, know, maybe, perhaps, and guess.

In terms of emotions, positive feelings were reported more often than negative ones, 5.62% of the time by the client, and 5.45% of the time by the counsellor. The skewing of the data on the client's data set could be interpreted due to filling in diary logs that were included in the text analysis. In the diary logs, the client often report about situations in which they want to drink, which is often mentioned with positive impact that drinking has on them (e.g. relaxation, being with friends). In case number two, the client often identified negative emotions, such as frustration, being mad, low energy, and exhaustion. In the content of email, she often reports similar emotions, along with guilt, low motivation, and feeling blue. When it comes to counsellor, he or she often addressed all the emotions the client mentioned, including supportive words and optimistic thoughts. In other cases, the difference between different prevalence in positive and negative emotions is more apparent.

Figure 6 present how positive and negative emotions emerged over the time of the correspondence.

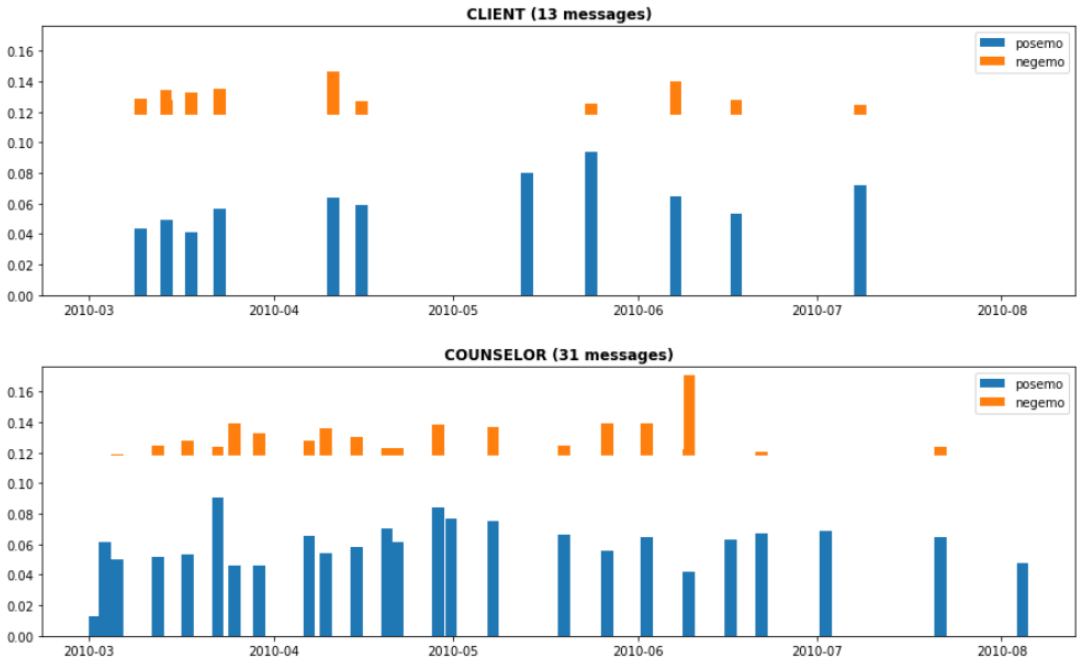



*Figure 6.* Positive and negative emotions over the time of correspondence

If the content of the emails is looked into in the times when the values reach their peak for the client between fifth and sixth month, it is observable that positive emotions were not represented in the content: "*I felt I had let myself down and the program and was feeling low in spite of everything I was trying to do."* On the other hand, positive emotions could be seen in sentences such as: *"Thank you for the suggestions on relaxing. […] I really love swimming and walking. I also love painting and have started to do paint".* In the case of counsellor, positive feelings that were mentioned include motivational speech, gratitude for participation and keeping up with good work. In the corresponding month, the counsellor simultaneously addressed client's fears and provided support for process, which is observable in following examples: *"Thank you for your recent message. I am sorry to hear that you have been feeling low and that you feel you have let yourself down this last weekend",* or "*The fact that you attended the surgery and spoke to the triage nurse about your difficulties is a great example of a healthy coping response – well done! It also demonstrates your continuing motivation and commitment to making changes.*

*Please remember how far you have come. A lapse does not mean total failure and one slip does not need to lead to a relapse to old behaviour patterns. There is no failure only feedback".* This portrays how both positive and negative emotions were represented within this timeframe.

If affective processes are differentiated into anger, sadness, and anxiety, the findings show that only sadness was consistently present, while anxiety was most often reported in the drinking diaries. Figure 7 and 8 portray how these subcategories of affective processes were distributed over the time.

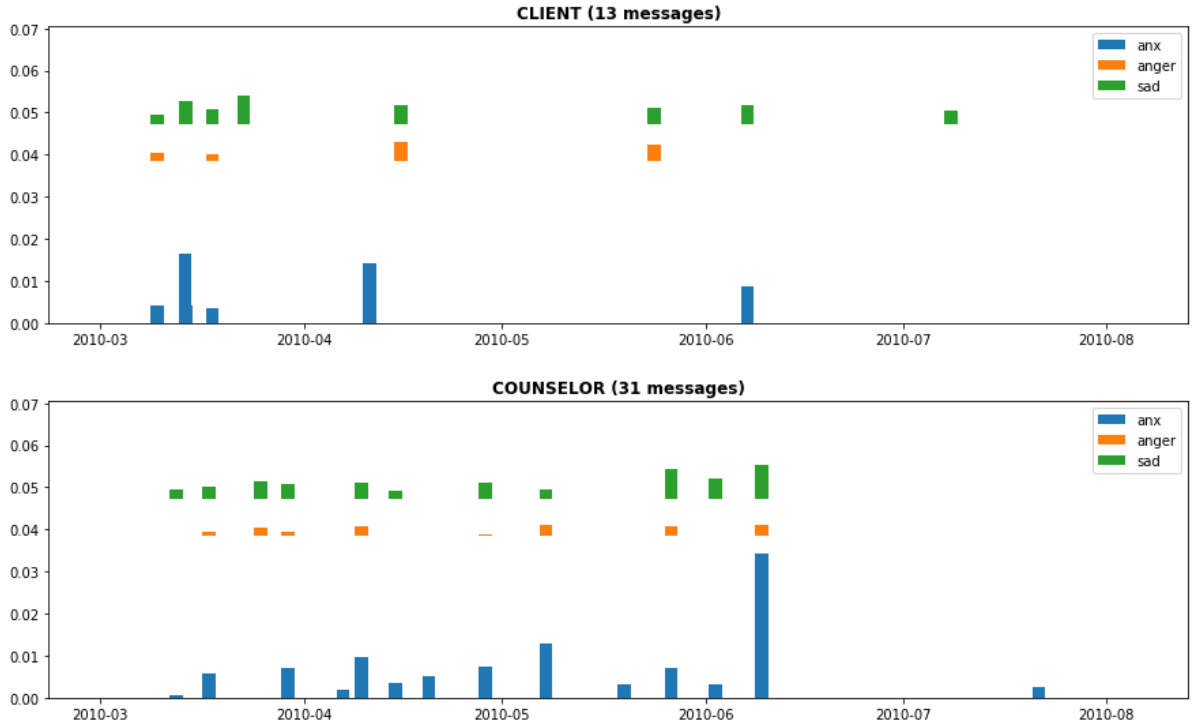



*Figure 7.* Anxiety, anger, and sadness for case two

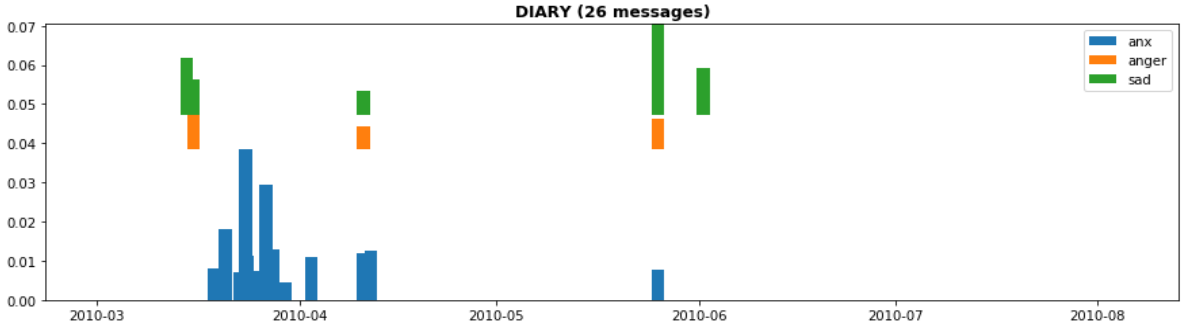



*Figure 8.* Anxiety, anger, and sadness in diaries

The prevalence of sadness and anxiety subcategories were apparent from the content of the emails and diaries. Client often referred to her anxiety, both regarding her feelings towards the change, or when she would use alcohol to alleviate these symptoms. Sadness was observable through emails regarding depressive episodes that client experienced.

LIWC allowed the researcher to observe different categories over the time within the context. For example, if dictionary terms from the LIWC Manual (Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007) were searched within the emails, it is observable how different categories could overlap in one sentence of an email. Table 8 contains an overview of examples found by using search terms in the data.

Table 8

*Manual LIWC search with email snippets*

| Processes | Client's Examples | Counsellor's Examples |
|---|---|---|
| Affect | "I felt defiance, anger and total disregard. I was drinking because I was scared. I knew that that I had exceeded my own limits I set" | "Appreciate your own efforts and the little successes." |
| | "I can relax more easily; I'm less bored; I'm less troubled by shaking, or feeling sick; I feel less emotional pain" | "Such negative emotions and thoughts are such a burden and would wear anyone down" |
| Cognitive | "That I am not going to be able to keep to this very rigid and structured cut down. I am thinking that my hackles are rising and that a mutinous attitude is setting in" | "Many people tend to think negative thoughts because they've learned to over the years , and they're used to it" |
| Social | "I agree with you that i need to parent myself but I found myself under immense pressure and felt myself dig my heels in and could not overcome this with any sort of rational thinking." | "You can also expect positive reactions from family, friends and acquaintances." |

This table portrays how overlapping of terms of LIWC categories is inevitable, and that interpretation without the context should not be taken at face value. If these three categories were divided in sub-categories, it is observable how different affective and cognitive processes were manifested in the data.

When Referential Activity (RA) is observed, as shown in Figure 9, the process of the intervention seems to be synchronized in increases of the DAAP values between the counsellor and the client. This indicates that within same timeframe, both client and counsellor were involved in one of the three processes Arousal, Symbolizing, and Reorganizing, which enrich and support reaching the positive outcomes in psychotherapy.

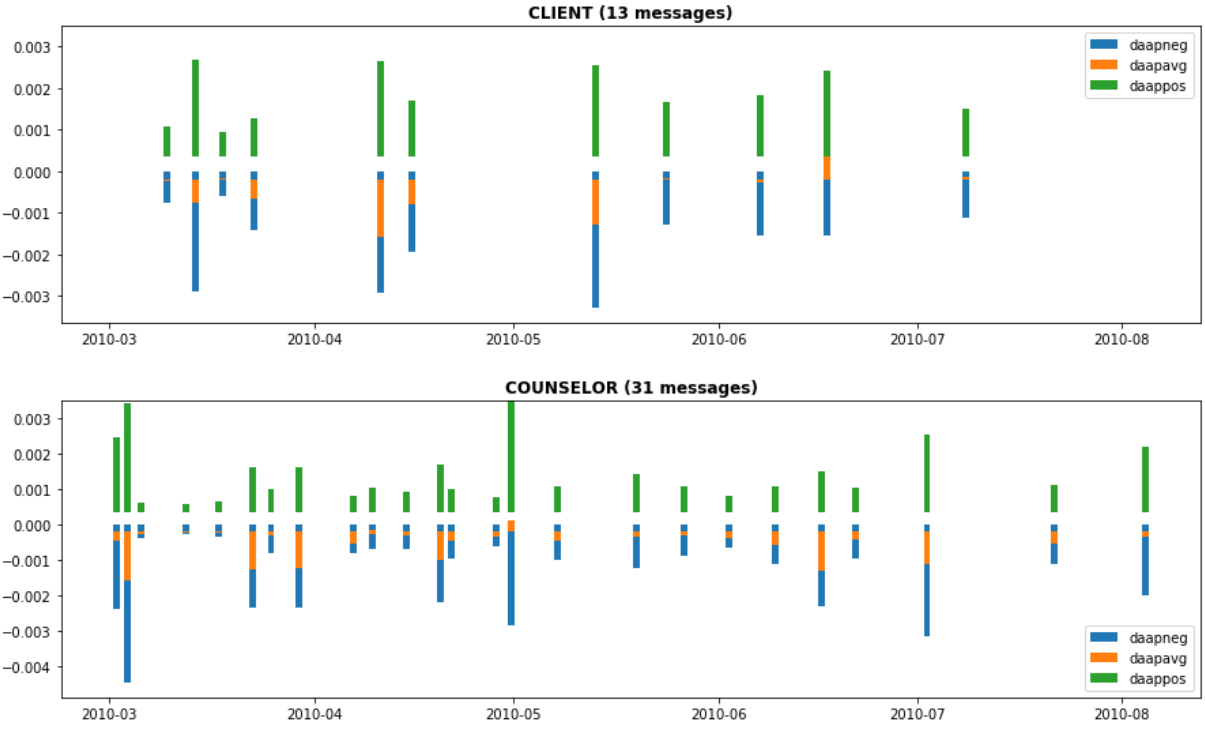



*Figure 9.* Referential Activity for case number two

The positive values of RA occur within the same timeframe with the expansion of cognitive and abstract language, as measured with LIWC. These results indicate that with simultaneous work, client and counsellor synchronized in their correspondence.

In case two, RA came in cycles, which is reflected in positive measures obtained via DAAP within the timeframe. For example, counsellor started the treatment using language with high RA in the first month (2010-03 on the graph), and client responded with high RA in the following emails from 2010-03 to 2010-04. In some timeframes, higher negative RA values occur within positive ones, indicating that not all of the content of the email have the same value in terms of describing experiences in a way that it reflects mental content with and observable emotional tone and feelings underneath.

If content of emails during high RA values is taken into account, the process behind the measures becomes clearer. For example, counsellor starts the process with higher RA in the beginning of the treatment referring to client's questionnaire about detecting advantages and disadvantages of drinking, exploring her habits and feelings with statements such as: *"You mentioned having moments when you feel down and suffer dark moods . You also said that you have felt suicidal once or twice. As I am concerned for your safety, I wanted to ask you a bit more about this."* After this, counsellor proceeds with more detailed inquiry about when, how, and why these feelings occur and if the client has significant relationship that could provide necessary support. The client engaged after this message, giving a detailed overview while answering these questions and providing additional information what she has tried in the past and how she tackles her negative thoughts when they occur. For example, client says: *"When it comes to my drinking sometimes I am not aware what makes me go the shop and buy alcohol, I just seem to be in the zone. I do not how to explain it otherwise. I will try to record exactly when*

*I feel like a drink , if I can be honest , in my brain I feel that this is a sort of technique for me to question whether I want the drink or not. I have tried this, when I have walked around my block, asking myself, do you really want to drink?"* In both of these examples, both client and counsellor adjusted the content of the emails in order to tackle client's negative feelings and thoughts.

When individual emails with high or low values are observed, RA activation is depicted like in Figure 10.

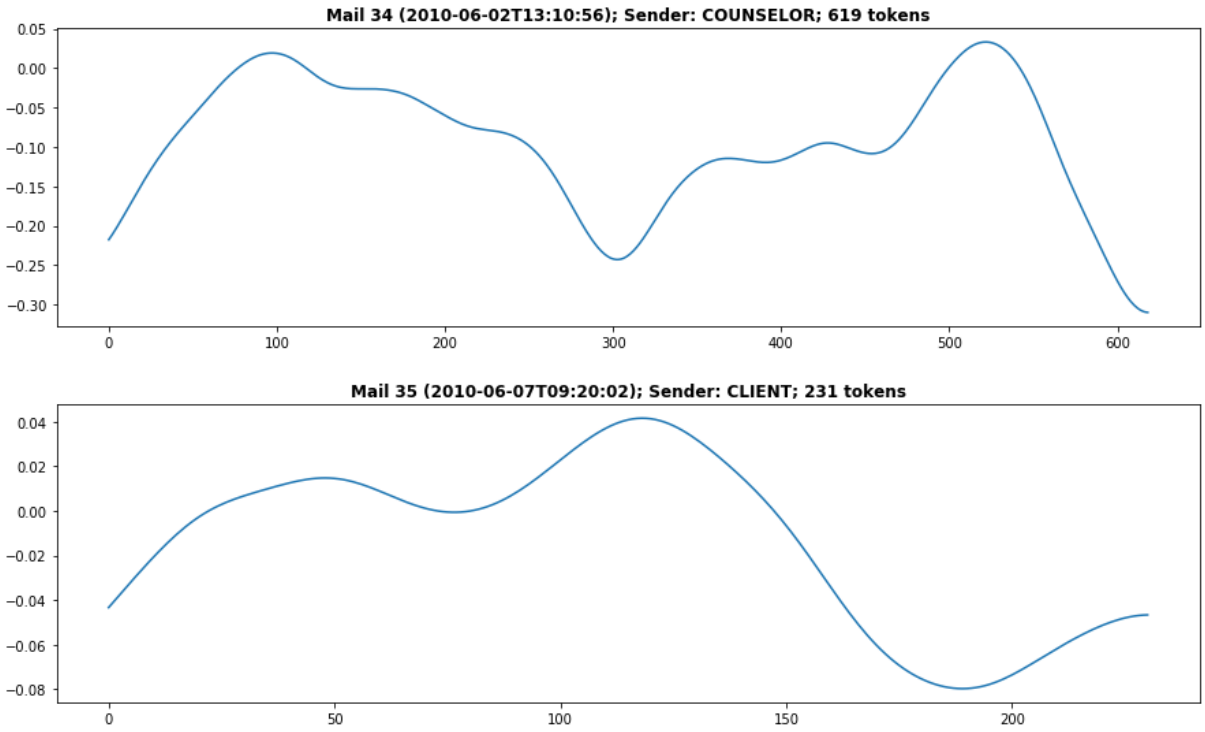



*Figure 10.* Referential Activity within given timeframe for case number two

In this correspondence, counselor starts the email expressing understanding for client's stressful situation at work, and emotional burden she is experiencing, offering alternative ways of thinking, but also suggesting reaching out for additional help. Counsellor also provides detailed information about how changes affect one's life and what feelings and thought could be useful or

aggravating for cutting down the drinking. . Therapist reassures the client to keep up with good work with words: *"As a participant in this therapy programme, you can give yourself a pat on the back. You've found the courage to really look at the problem, and you're even actively working on changing your drinking habits. Those are very positive actions! PERSON do not be too hard on yourself. Appreciate your own efforts and the little successes. On days when things don 't work , it can help to see it as a learning experience : it didn't 't go as I wanted it to ; so what can I learn from it ? What am I satisfied with? Might I be happy with a little less, just once".* The decrease in RA activity for the counsellor happens in the middle of the email, where she or he forwarded a snippet from an article which is related to the context. These piece of emails shows how words that further exploration of information given by the client herself tap into emotions behind higher RA scores.

Client's response includes her reaction to her inability to cut down drinking and she reports feeling guilty, low, and how she searched for additional help, along with agreeing with the therapist how she should address the problems regarding immerse pressure that she has been putting on herself. This is reflected in sentences, such as: "*Over the past NUM weeks since I have not been able to cut down my intake to our suggested levels I have felt a failure. What came up for me was painful memories of childhood of not being good enough. Feelings of shame, guilt. Yes I do know I am trying and I was ready to give up the online therapy but saw the nurse again who asked me to continue but believes I do need additional support."* As can be seen in the Figure 10, the decrease in client's RA activity occurred after two thirds of the content (after 150 words), in which the client states: "*I agree with you that I need to parent myself but I found myself under immense pressure and felt myself dig my heels in and could not overcome this with*

*any sort of rational thinking"*. In this email snippet, client uses expressions and gives a less in-depth information about how she feels.

Bearing in mind that this example encompasses both high and low RA, the evolution of the emotion-abstract language becomes apparent, demonstrating how words that are more cognitively and affectively charged and related to client's current problems affect the values in total. In other cases with low RA activity, the exchange was usually based on technical information, or arranging a substitute therapist while the assigned therapist is on vacation.

Figure 11 shows the last emails exchanged before the client dropped out of the study.

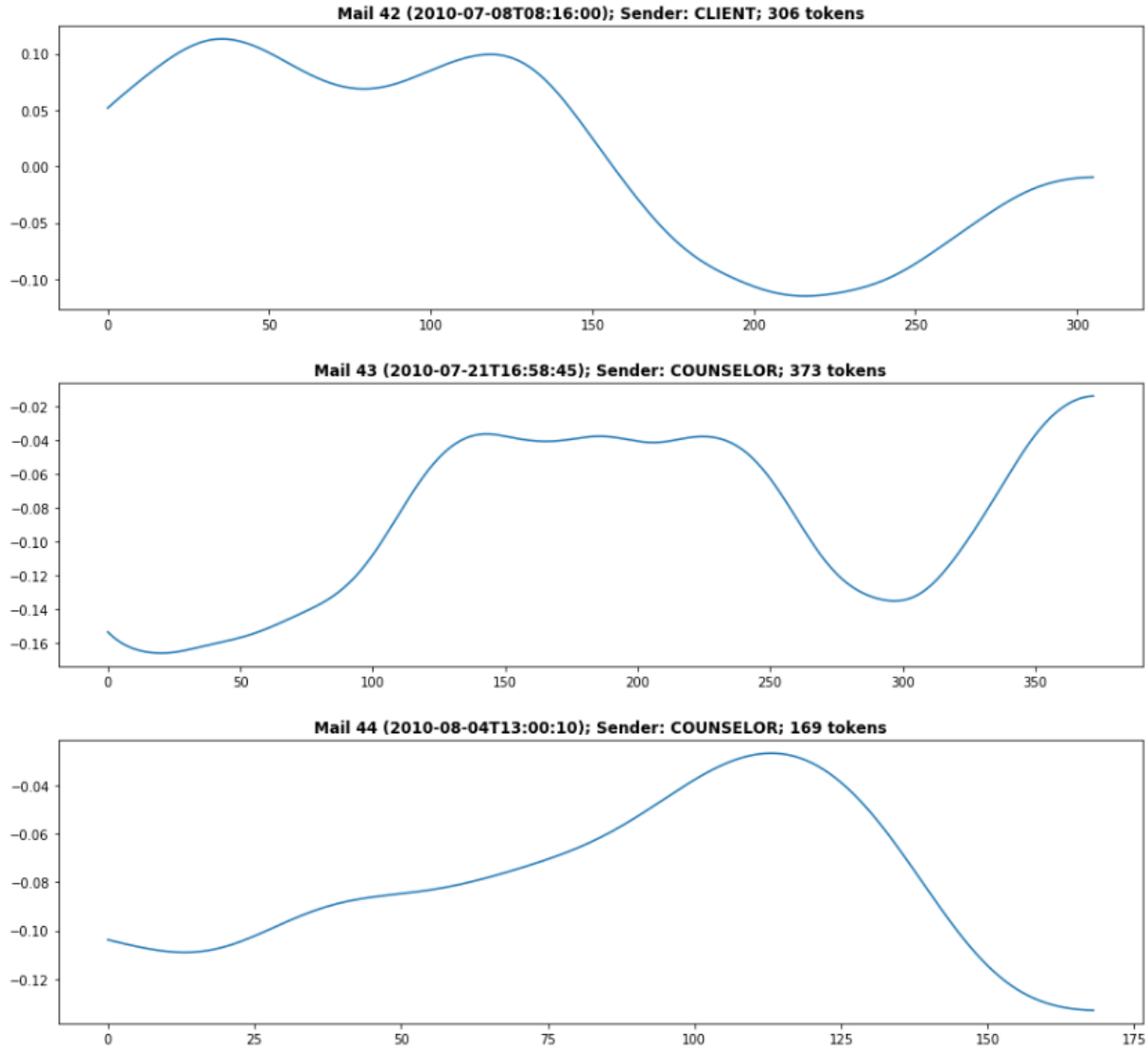



*Figure 11.* Referential Activity at the end of the treatment for client number two

The last email that client had sent before she dropped out of the programme included information about more positive feelings towards the programme and how the client reached out for help outside the web-intervention. The decrease in RA activity occurred when she stated that she will explore her drinking habits further, but that she failed before. Once she reflects how she plans to tackle this problem in the future: *"I'll just break it anyway at this stage. I have not managed to do alcohol free days and I have also discussed this with my counsellor. Knowing how I am, I believe I have to take the approach of trying to reduce the daily consumption, then week*

*by week, to try to consolidate some alcohol free days."*, the RA increases again. The counsellor made an attempt to check on the client and see if she is still interested in continuing with the programme, expressing understanding for her need to talk to someone else, but also worry that reaching out for extra help might be burdening at that moment. Additionally, the therapist offer to help the client with logbook, explaining why is it important for the client to maintain filling them in. In the last email, the counsellor gives standard information about the client's inactivity and assumes that she is no longer interested in participating. When RA activity of this email increases, the counsellor leaves an open invitation for the client to join again.

## Discussion

The main aim of this study was to demonstrate how use of emotion-abstract language is manifested in the sample of the web-based "Look at your drinking" intervention, and to examine if different text mining techniques underpin further investigation of using such language in virtual setting. The provided results signified how preselected LIWC categories and Referential Activity (RA) measured with DAAP emerge in the whole sample, and within one single case. The mixed method approach allowed the researcher to interpret obtained findings in the context, as well as to identify and predict limitations and potentials of using such tools in the future research.

### Identification of relevant LIWC categories

The results gathered using LIWC text mining technique have highlighted primarily the differences of the number of tokens (words) used by the clients and the therapists. On average, the pool of used tokens was much higher in counsellors, than in clients. Other findings have shown that categories social, cognitive, and affective were the most dominant one across the sample. If the content of the emails was taken into account, these categories correspond to the

problems, situations, and feelings that are addressed most often in the text, indicating that emotion-abstract language was used for gaining deeper understanding and working through these topics. However, the increases in social category within both senders could be skewed due to the generic emails and the fact that diary logs were included in analysis. Due to the nature of the intervention, which often addresses the social factors relevant for drinking, and that diary logs were included in the analysis, increase in social category was expected.

Taking content into consideration, the distribution of the percentages of positive and negative emotions and their subcategories was not expected. It seemed that LIWC linguistic markers for affective words do not correspond to the content of the emails. The unanticipated differences in the emotions could be explained with including diaries in the analysis. The emotions in the diaries indicate that alcohol consumption is often in contingency with social interactions or alleviating effects of drinking. The diaries showed some discrepancies between reported emotions and thoughts when compared to the content of the emails. Understanding how emotions evolve during the treatment, but also in diary data could provide valuable insights on how clients verbalize their affective processes. Althoff, Clark, and Leskovec (2016) used positive and negative emotion words to explore how these linguistic aspects correlate with mental health counselling success outcomes, showing that changes towards more positive word usage predicts better outcomes. Mergenthaler (2008) found that negative emotions usually occur in the beginning of the therapeutic process, whereas positive ones emerge later. Additionally, Fan and al. (2019) reported that affect labelling, or putting feelings into words, in online surrounding can diminish effects of positive and negative emotions.

Cognitive processes were dominantly used by both senders. Both clients and counsellors used these subcategories with greater frequency, with higher values measured among

counsellors. Since vast majority of generic emails encompassed thinking patterns, these results were foreseeable. Personal issues did not seem to have strong prevalence across the content, in spite of the expectations that they will be addressed often, due to the fact that client often report how their excessive drinking affects their work and free time. The observation that the most prominent categories increase at the same time for both counsellors and clients could indicate attuning and adjusting of the used language. Some attempts to establish these overlapping as a Language Style Synchrony has already been taken, showing that therapeutic dyad can be observed in finding matching language styles using LIWC (Lord et al., 2015). This kind of matching is in accord with Hölzer et al. (2010) findings that accommodating language is more frequently reported in successful therapies.

**Referential Activity in virtual setting**

The results gathered using DAAP programme showed that the changes in terms of Referential Activity (RA) are observable in online correspondence. Bearing in mind that most of the studies using DAAP and WRAD dictionary are conducted on transcripts of the psychotherapeutic sessions, and that values range from -1 to +1, it is not possible to make unequivocal interpretation if these values that are derived on the whole correspondence within several months, and by nature contain different type of text is low or high. Additionally, this study is focused on online intervention, and DAAP has not been employed in this setting before. Usually, therapeutic sessions contain between 5000 and 7000 words in total (Murphy, 2015). In this study, average poll of words sent by clients was 39370 and 153688 for counsellors. Even though the data set in this research is noticeably larger, the text does not encompass one session, but the whole process in digital surrounding.

Nonetheless, the DAAP managed to identify via WRAD the psycholinguistic style of Arousal, Symbolizing and Reorganizing phases, as shown in previous studies (Murphy, 2015). When the content of the emails was taken into account, the RA activity coincides with problem solving situations, especially via tapping into affective and cognitive processes behind it, as shown in case analysis. Lower RA activity usually grasped the language regarding technical issues with the website, arranging substitute counsellor, or discussing vacation dates.

Referential Activity emerged in sequences across the whole sample, depicting how emotion-abstract language was utilized for exploration of underlying reasons for excessive drinking and emotions that clients experience during the intervention.

Connecting patterns could indicate presence of good therapeutic process (Verde, Sarracino, & Vigorelli, 2012). Mergenthaler (1996, 2008, 2017) has already shown how crucial events of therapy can be detected with identifying emotion-abstract patterns using Therapeutic Cycle Model (TCM). Such cycles are understood as emotion-abstraction patterns that could be observed either through connecting blocks in graphs, or block that follow one another (Mergenthaler, 2008). Both RA and TCM have showed promising results in connected patterns between the therapist and the client (Verde et al., 2012), indicating that assumption of clients and the counsellors between will be activated in problem solving situations (Mergenthaler, 2008).

**Overlapping of LIWC and DAAP findings**

LIWC and DAAP values increased around the same dates, indicating that predetermined categories related to emotion-abstract language and RA could be interpreted in a way that using words related to emotions, cognition, and describing the social factors of drinking habits are valuable to the online treatment if they are conveyed in a way to correspond to three phases behind the RA. Since these construct measure the same linguistic markers, the contingency of the

findings is not unexpected. However, the findings of this study portray how mere identification without context using LIWC could affect interpretation of the data. Referential activity and correlation with LIWC categories could bridge differences of these techniques. Gandino et al. (2017) already addressed the necessity of employing more detailed analysis in addition to LIWC, highlighting limitations of word-by-word count techniques and offering DAAP as a more elaborate analysis.

Analysis conducted on one case from the sample showed how RA occurred in written correspondence, allowing the researcher and readers to see how emotion-abstract language was utilized in the intervention. With mutual engagement on problem solving and exploring cognition and emotion, both senders managed to tap into reasons behind client's excessive drinking, which is reflected in increases in RA activities in both senders. These results are demonstrated in overall measures, but also when observed within the activity of a single emails. RA values emerged sequentially and LIWC categories seemed to follow the similar trend. LIWC analysis has shown that cognitive and affective language has been used by both clients and counsellors, arguing that attuning of the language occurred during the treatment. In eight out of nine cases, increase in one category within client is followed by the counsellors' increase within the same category, and vice versa. In terms of LIWC, the results advocate for gaining deeper understanding of the dominant categories.

Since RA findings correspond with increase in LIWC categories, further investigation of how these two text mining techniques could contribute to each other or how similar patterns could be identified is encouraged.

**Further research of emotion-abstract language use in virtual setting**

The findings obtained via mixed method approach in this study indicate that emotion-abstract language has potential in future web-deliverable programmes. Contemporary research that vastly implements MI and CBT approach in their online interventions is already well documented (Mihalcea et al., 2014; Postel, De Haan, et al., 2010; Tanana et al., 2017). The findings of this study exhibit how values obtained using different techniques can reflect the actual content of the provided text, highlighting the importance of finding a way to grasp the emotional and cognitive processes behind the language or between the lines.

Taken all of the presented results into consideration, existing cycles of RA and observable categories that were identified could have important implications for further research. As Lord, Baer, and Atkins (2015) suggested in their study, indicators of language style synchrony or matching are reflected in empathy ratings between the therapist and the client. Nonetheless, Mergenthaler (2008) findings indicate that counsellor and client have "resonating minds" and that during the psychotherapeutic process they learn about each other's preferred style of communication and interaction, which results in adjusting their language in favour of positive outcomes of the treatment. Using LIWC Mihalcea et al. (2014) showed that using syntactic patterns could increase the prediction of the counsellors' behaviour, implying that using congruent language could lead to better and clearer outcomes. In terms of innovative measures, linguistic analyses focused on the language of the therapist and the client throughout the session provide the means to identify structures inherently given in the natural language (Halfon, 2017).

Many studies justify the assumption that client's and counsellor's language are attuning to each other and that this mechanism could be crucial for therapeutic outcomes. The results regarding such language obtained used LIWC, which has already exhibited its ability to

determine psychological processes in online derived text, and DAAP, which has not previously used in virtual setting, demonstrate that further exploration of emotion and cognitive processes in digital surrounding is highly encouraged.

**Limitations of the study**

Limitations of this study are reflected in several areas. Firstly, in order to make conclusions regarding detecting emotion and abstract patterns, the study needs to be conducted on a bigger sample. Secondly, careful analysis of the data should be included. This entails in-depth examination of generic messages that are sent by the counsellors, along with measuring therapeutic outcomes of the treatment.

The results are presented in percentages relative to the overall token use, which affects the interpretation. Since scaling these results was impossible, the results from clients and counsellor could not be directly comparable. The potential skewing of the data and observed pattern in this study might be a by-product of replying to the emails or simply forwarding them. Since correspondence between the participants and counsellors is followed by assignments and questionnaires, the observed patterns could be identified due to responding to materials that were previously provided. The observed difference between numbers is expected, not only due to way the values were calculated, but also due to the fact that counsellors sent generic messages that include questionnaires and tailored information about the indicated problems of the client more frequently. The generic messages sent by counsellors are advised to be either excluded, or labelled as such, in order to prevent interference with the data.

Additionally, including therapeutic outcomes would help concluding if the matching in emotion-abstract language could predict positive outcomes in online treatments.

LIWC categories are identified based on a dictionary which reflects the words that are relevant to the category. In some of the cases of the study, spirituality was mentioned multiple times, but category related to religion remained 0%. In the analysed case, praying and God were mentioned several times, but the same category remained without detected tokens. Additionally, the provided email snippets imply that detection of the categories without additional analysis cannot grasp the complete depth of the exchanged thoughts, feelings, and processes behind the correspondence.

For this study, only one DAAP dictionary (WRAD) was available. In order to detect three different phases that comprise Referential Activity, utilizing REF dictionary (Murphy, 2015) is necessary.

## Conclusion

"Human cognition, emotion, and behaviour lie at the heart of the therapeutic process" (McCarthy, Caputi, & Grenyer, 2017, p. 383)

Both LIWC and DAAP have managed to identify emotion-abstract language within the email correspondence in the Web-based intervention "Look at your drinking". The predetermined LIWC categories cognitive processes, perception, social, affect, biological, and personal concerns have emerged in contingency between counsellors' and clients' exchanged emails. Positive emotions were more prevalent the negative ones, especially in diaries. Both positive and negative emotions have important role in psychotherapy - studies suggest that negative emotions facilitates the presentations of problematic materials, whereas positive ones are utilized in problem solving situations (McCarthy et al., 2017). This study reported how chosen LIWC categories emerge within the emails of participants and counsellors, showing that without context their interpretation might differ from the conveyed emotions and thoughts.

Using DAAP technique, the changes within the therapy have emerged in sequences, usually followed by similar RA values within the messages. Since RA has shown to be measurable in a web-based intervention, more research on how this could improve the relationship between the clients and counsellor is encouraged. Knowing that DAAP has been already used for identifying synchrony and positive outcomes in therapy (Bucci & Maskit, 2007; Bucci & Murphy, 2015; Murphy, 2015), its utility in digital surrounding is yet to be discovered. Previous research has shown that gains in therapy that are identified earlier in the session are in connection with emotional and cognitive change, no matter of the preferred school of psychotherapy (CBT, psychodynamic, or other) (McCarthy et al., 2017).

In addition to this, LIWC and DAAP showed overlapping results in terms of the preselected categories and higher Referential Activity.

These results indicate that further research should look into language that could be collected within the interventions, in order to improve alliance, empathy, rapport, and mutual understanding. Since most successful online interventions so far fall under domain of excessive drinking and addictions, these treatments could be the important starting point for improvements.

The inevitable challenge in online interventions lies in establishing profound relationship in a changed context with limited space for detecting cues such as intonation, (non)verbal language, and overall sense of the therapeutic situation. Bearing in mind that the therapeutic process encompasses changes over the time, rather than random distributions and immediate occurrences through the treatment (McCarthy et al., 2011), more detailed insights into how processes analysed in this study contribute to positive outcomes is encouraged.

In light of the collected data, utilising emotion-abstract language is relevant for further development of delivering such services online.

References

Althoff, T., Clark, K., & Leskovec, J. (2016). Large-scale Analysis of Counseling Conversations : An Application of Natural Language Processing to Mental Health, *4*, 463–476.

*American Psychological Association*. (2019, January 27). Retrieved from Understanding Psychotherapy: https://www.apa.org/helpcenter/understanding-psychotherapy

Andry, F., Freeman, L., Gillson, J., Kienitz, J., Lee, M., Naval, G., … Mateo, S. (2008). Highly-Interactive and User-Friendly Web Application for People with Diabetes. *Medical Informatics*, 118–120. Retrieved from http://elifesensor.com/healthcom2008/ICW_IEEE_HEALTHCOM08_FINAL.pdf

Barrett, L. F. (2004). Feelings or Words ? Understanding the Content in Self-Report Ratings of Experienced Emotion, *87*(2), 266–281. https://doi.org/10.1037/0022-3514.87.2.266

Bellot, P., Bonnefoy, L., Bouvier, V., & Duvert, F. (2013). *Chapter 1 Large Scale Text Mining Approaches for Information Retrieval and Extraction*. https://doi.org/10.1007/978-3-319-01866-9

Bruner, J. (1991). The Narrative Construction of Reality, *18*(1), 1–21.

Bucci, W. (2013). The Referential Process as a Common Factor Across Treatment Modalities, *16*(1), 16–23. https://doi.org/10.7411/RP.2013.003

Bucci, W., & Maskit, B. (2006). A Weighted Referential Activity Dictionary . In J. G. Shanahan, Y. Qu, & J. Wiebe, *Computing Attitude and Affect in Text: Theory and Applications* (pp. 49-60). Springer .

Bucci, W., & Maskit, B. (2007). Beneath the surface of the therapeutic interaction: The psychoanalytic method in modern dress. *Journal of the American Psychoanalytic Association*, *55*(4), 1355–1397. Retrieved from https://doi.org/10.1177/00030651070550041 2

Bucci, W., Maskit, B., & Murphy, S. (2016). Connecting emotions and words : the referential process. *Phenomenology and the Cognitive Sciences*, (February 2015), 359–383. https://doi.org/10.1007/s11097-015-9417-z

Bucci, W., & Murphy, S. M. (2015). Putting Feelings into Words : Cross-Linguistic Markers of the Referential Process, 80–88.

Canfield, M. L., Walker, W. R., & Brown, L. G. (1991). Contingency Interaction Analysis in Psychotherapy, (1), 58–66.

Caragea, C. (2017). Identifying Empathetic Messages in Online Health Communities, 246–251.

Clarke, G., Eubanks, D., Reid, E., Kelleher, C., Connor, E. O., Debar, L. L., … Clarke, G. (2005). Overcoming Depression on the Internet ( ODIN ) ( 2 ): A Randomized Trial of a Self-Help Depression Skills Program With Reminders Corresponding Author :, *7*(2), 1–12. https://doi.org/10.2196/jmir.7.2.e16

Copeland, J., Ph, D., & Martin, G. (2004). Web-based interventions for substance use disorders : A qualitative review, *26*, 109–116. https://doi.org/10.1016/S0740-5472(03)00165-X

Cunningham, J. A., & Breslin, F. C. (2004). Only one in three people with alcohol abuse or dependence ever seek treatment, *29*, 221–223. https://doi.org/10.1016/S0306-4603(03)00077-7

Etter, J., & Geneva, C.-. (2005). Comparing the Efficacy of Two Internet-Based , Computer-Tailored Smoking Cessation Programs : A Randomized Trial Corresponding Author :, *7*. https://doi.org/10.2196/jmir.7.1.e2

Fan, R., Varol, O., Varamesh, A., Barron, A., Leemput, I. A. Van De, Scheffer, M., & Bollen, J. (2019). The minute-scale dynamics of online emotions reveal the effects of affect labeling. *Nature Human Behaviour*, *3*(January). https://doi.org/10.1038/s41562-018-0490-5

Fast, E., Chen, B., & Bernstein, M. S. (2016). Empath : Understanding Topic Signals in Large-Scale Text.

Gandino, G., Fini, G. Di, Bernaudo, A., Paltrinieri, M., Castiglioni, M., & Veglia, F. (2017). The impact of perinatal loss in maternity units : A psycholinguistic analysis of health professionals ' reactions. https://doi.org/10.1177/1359105317727841

Gibson, J., Malandrakis, N., Romero, F., Atkins, D. C., & Narayanan, S. (2015). Predicting Therapist Empathy in Motivational Interviews using Language Features Inspired by Psycholinguistic Norms. *In Proceedings of Interspeech*, 2–6.

Gibson, J., Xiao, B., Imel, Z. E., Atkins, D. C., Georgiou, P., & Narayanan, S. (2016). A Deep Learning Approach to Modeling Empathy in Addiction Counseling, (76), 1447–1451.

Halfon, S. (2017). An Empirical Study of Verb Use as Indicator of Emotional Access in Therapeutic Discourse, *34*(1), 35–49.

Haug, S., Strauss, B., Gallas, C., Kordy, H., Haug, S., Strauss, B., … Kordy, H. (2008). New prospects for process research in group therapy : Text-based process variables in psychotherapeutic Internet chat groups, *3307*. https://doi.org/10.1080/10503300701368008

Hölzer, M., Mergenthaler, E., Pokorny, D., Kächele, H., Hölzer, M., Mergenthaler, E., … Luborsky, L. (2010). Vocabulary Measures for the Evaluation of Therapy Outcome : Re-Studying Transcripts From the Penn Psychotherapy Project EVALUATION OF THERAPY OUTCOME : PEN ” PSYCHOTHERAPY PROJECT, *3307*. https://doi.org/10.1080/10503309612331331618

Howes, C., & Purver, M. (2012). Linguistic Indicators of Severity and Progress in Online Text-based Therapy for Depression ∗, (611733).

Humphreys, K., & Klaw, E. (2001). Can Targeting Nondependent Problem Drinkers and Providing Internet-Based Services Expand Access to Assistance for Alcohol Problems ? A Studi of the Moderation Management Self-Help / Mutual Aid Organization *, (August 2001). https://doi.org/10.15288/jsa.2001.62.528

Imel, Z. E., Steyvers, M., & Atkins, D. C. (2016). Computational Psychotherapy Research: Scaling up the evaluation of patient-provider Interactions. *Psychotherapy*, *52*(1), 19–30. https://doi.org/10.1037/a0036841.Computational

Kosinski, M., Wang, Y., Lakkaraju, H., & Leskovec, J. (2016). Mining Big Data to Extract Patterns and Predict Real-Life Outcomes, *21*(4), 493–506.

Lambiotte, B. R., & Kosinski, M. (2014). Tracking the Digital Footprints of Personality, *102*(12), 1934–1939.

Lord, S. P., Baer, J., & Atkins, D. C. (2015). More Than Reflections : Empathy in Motivational Interviewing Includes Language Style Synchrony Between Therapist and Client. *Behavior Therapy*, *46*(3), 296–303. https://doi.org/10.1016/j.beth.2014.11.002

Lord, S. P., Sheng, E., Imel, Z. E., Baer, J., & Atkins, D. C. (2016). includes language style synchrony between therapist and client, *46*(3), 296–303. https://doi.org/10.1016/j.beth.2014.11.002.More

Low, K. G., Charanasomboon, S., Lesser, J., Reinhalter, K., Martin, R., Jones, H., … Martin, R. (2006). Effectiveness of a Computer-Based Interactive Eating Disorders Prevention Program at Long-Term Effectiveness of a Computer-Based Interactive, *0266*. https://doi.org/10.1080/10640260500403816

Lundahl, B., & Burke, B. L. (2009). The Effectiveness and Applicability of Motivational Interviewing : A Practice-Friendly Review of Four Meta-Analyses, *65*(11), 1232–1246. https://doi.org/10.1002/jclp

McCarthy, K. L. (2015). Emotional and cognitive processes in psychotherapy for complex mental health disorders : Linguistic analyses of significant therapist-patient interactions.

Mccarthy, K. L., Caputi, P., & Grenyer, B. F. S. (2017). Significant change events in psychodynamic psychotherapy : Is cognition or emotion more important ?, 377–388. https://doi.org/10.1111/papt.12116

McCarthy, K. L., Mergenthaler, E., Schneider, S., Grenyer, B. F. S., Mccarthy, K. L., Mergenthaler, E., … Grenyer, B. F. S. (2011). Psychodynamic change in psychotherapy: Cycles of patient–therapist linguistic interactions and interventions, *3307*. https://doi.org/10.1080/10503307.2011.615070

McKenna, K. Y. A., & Green, A. S. (2002). Virtual Group Dynamics, *6*(1), 116–127. https://doi.org/10.1037//1089-2699.6.1.116

Mergenthaler, E. (1996). Emotion-Abstraction Patterns in Verbatim Protocols : A New Way of Describing Psychotherapeutic Processes, *64*(6), 1306–1315.

Mergenthaler, E. (2008). Resonating minds : A school-independent theoretical conception and its empirical application to psychotherapeutic processes, *3307*. https://doi.org/10.1080/10503300701883741

Mergenthaler, E. (2017). The therapeutic cycles model in psychotherapy research : Theory , measurement and clinical application, *9338*(November). https://doi.org/10.1016/S0924-9338(97)80391-4

Mihalcea, R., Resnicow, K., Singh, S., An, L., Goggin, K. J., Catley, D., … Services, H. (2014). Predicting Counselor Behaviors in Motivational Interviewing Encounters.

Miller, W. R., & Rose, G. S. (2010). Toward a Theory of Motivational Interviewing, *64*(6), 527–537. https://doi.org/10.1037/a0016830.Toward

Mohammad, S. M. (2015). Sentiment Analysis : Detecting Valence , Emotions , and Other Affectual States from Text.

Murphy, S. (2015a). A Computer Program for Tracking the Evolution of a Psychotherapy Treatment, (1), 134–145.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The Development and Psychometric Properties of LIWC2015

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). The Development and Psychometric Properties of LIWC2007 The University of Texas at Austin, 1–22.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words , Our Selves. https://doi.org/10.1146/annurev.psych.54.101601.145041

Pennebaker, J. W., & Seagal, J. D. (1999). Forming a Story : The Health Benefits of Narrative, *55*(10), 1243–1254.

Postel, M. G. (2011). *Well Connected.Web-based treatment for problem drinkers*. Oldenzaal, the Netherlands: Nijmegen Institue for Scientist-Practitioners in Addiction, Radboud University Nijmegen.

Postel, M. G., De Haan, H. A., Ter Huurne, E. D., Becker, E. S., & De Jong, C. A. J. (2010). Effectiveness of a web-based intervention for problem drinkers and reasons for dropout: Randomized controlled trial. *Journal of Medical Internet Research*, *12*(4), 1–12. https://doi.org/10.2196/jmir.1642

Postel, M. G., De Jong, C. A. J., & De Haan, H. A. (2005). Does E-Therapy for Problem Drinking Reach Hidden Populations? *The American Journal of Psychiatry*, (162 (12)), 2393.

Postel, M. G., Haan, H. A. De, Huurne, E. D., Becker, E. S., & Jong, C. A. J. De. (2010). Effectiveness of a Web-based Intervention for Problem Drinkers and Reasons for Dropout : Randomized Controlled Trial Corresponding Author : Results :, *12*, 1–12. https://doi.org/10.2196/jmir.1642

Postel, M. G., Haan, H. A. De, & Jong, C. A. J. De. (2010). Evaluation of an E-Therapy Program for Problem Drinkers : A Pilot Study Treatment : E-Therapy Evaluation of an E-Therapy Program for Problem Drinkers : A Pilot Study, *6084*(May). https://doi.org/10.3109/10826084.2010.481701

Postel, M. G., Haan, M. S. H. A. De, Jong, M. D. C. A. J. De, & Ph, D. (2008). E-Therapy for Mental Health Problems :, *14*(7), 707–714.

Rees, C. S., & Stone, S. (2005). Therapeutic Alliance in Face-to-Face Versus Videoconferenced Psychotherapy, *36*(6), 649–653. https://doi.org/10.1037/0735-7028.36.6.649

Richards, D., Richardson, T., Timulak, L., & Mcelvaney, J. (2015). The ef fi cacy of internet-delivered treatment for generalized anxiety disorder : A systematic review and meta-analysis. *INVENT*, *2*(3), 272–282. https://doi.org/10.1016/j.invent.2015.07.003

Rodda, S. N., Booth, N., Vacaru, M., Knaebe, B., & Hodgins, D. C. (2018). Behaviour change strategies for internet, pornography and gaming addiction: A taxonomy and content analysis of professional and consumer websites. *Computers in Human Behavior*, *84*, 467–476. https://doi.org/10.1016/j.chb.2018.03.021

Roesler, C. (2017). Tele-analysis : the use of media technology in psychotherapy and its impact on the therapeutic relationship, 372–394. https://doi.org/10.1111/1468-5922.12317

Roskam, R. V. (2013). An Explorative Study for Therapist Treatment Fidelity in an Asynchronous Web - Based Treatment.

Salvatore, S., Gennaro, A., Auletta, A. F., Nitti, M., Salvatore, S., Gennaro, A., … Nitti, M. (2012). Automated method of content analysis : A device for psychotherapy process research, *3307*. https://doi.org/10.1080/10503307.2011.647930

Sassaroli, S., Brambilla, R., Cislaghi, E., Colombo, R., & Centorame, F. (2014). Emotion-Abstraction Patterns and Cognitive Interventions in a Single Case of Standard Cognitive-Behavioral Therapy, (April 2015). https://doi.org/10.7411/RP.2014.020

Sinha, A., Porter, T., Hons, B. A., & Wilson, A. (2018). The Use of Online Health Forums by Patients With Chronic Cough : Qualitative Study Corresponding Author :, *20*. https://doi.org/10.2196/jmir.7975

*Standard drink*. (n.d.). Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Standard_drink

Strecher, V. J., Shiffman, S., & West, R. (2005). Randomized controlled trial of a web-based computer-tailored smoking cessation program as a supplement to nicotine patch therapy. *Addiction*, *100*(5), 682–688. https://doi.org/10.1111/j.1360-0443.2005.01093.x

Tanana, M., Hallgren, K. A., Imel, Z. E., Atkins, D. C., Srikumar, V., States, U., … States, U. (2017). A Comparison of Natural Language Processing Methods for Automated Coding of Motivational Interviewing, 43–50. https://doi.org/10.1016/j.jsat.2016.01.006.A

Tanana, M., Hallgren, K., Imel, Z., Atkins, D., Smyth, P., & Srikumar, V. (2015). Recursive Neural Networks for Coding Therapist and Patient Behavior in Motivational Interviewing, 71–79.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words : LIWC and Computerized Text Analysis Methods. https://doi.org/10.1177/0261927X09351676

*The Discourse Attributes Analysis Program DAAP* . (2019, February 11). Retrieved from The Referential Process: http://www.thereferentialprocess.org/the-discourse-attributes-analysis-program-daap

Tjong, E., Sang, K., Vries, B. De, Smink, W., Veldkamp, B., Westerhof, G., & Sools, A. (2019). De-identification of Dutch Medical Text, 1–4.

Truong, K. P., Westerhof, G. J., Lamers, S. M. A., Jong, F. De, & Sools, A. (2013). Emotional Expression in Oral History Narratives : Comparing Results of Automated Verbal and Nonverbal Analyses, *231287*(231287), 310–314. https://doi.org/DOI: 10.4230/OASIcs.CMN.2013.310

Ubhi, H. K., Kotz, D., Michie, S., van Schayck, O. C. P., Sheard, D., Selladurai, A., & West, R. (2016). Comparative analysis of smoking cessation smartphone applications available in 2012 versus 2014. *Addictive Behaviors*, *58*, 175–181. https://doi.org/10.1016/j.addbeh.2016.02.026

*Unit of alcohol*. (n.d.). Retrieved from Wikipedia : https://en.wikipedia.org/wiki/Unit_of_alcohol

Verde, R. Lo, Sarracino, D., & Vigorelli, M. (2012). Therapeutic Cycles and Referential Activity in the Analysis of the Therapeutic Process, *15*(1), 22–31. https://doi.org/10.7411/RP.2012.003

Westerhof, G. J., Lamers, S. M. A., & Jong, F. De. (2014). Towards modeling expressed emotions in oral history interviews : Using verbal and nonverbal signals to track personal narratives, *29*(4). https://doi.org/10.1093/llc/fqu041

Xiao, B., Imel, Z. E., Georgiou, P., Atkins, D. C., & Narayanan, S. S. (2017). Computational Analysis and Simulation of Empathic Behaviors: A Survey of Empathy Modeling with Behavioral Signal Processing Framework, *18*(5), 1–18. https://doi.org/10.1007/s11920-016-0682-5.Computational

## Footnotes

[1] eHealth and telehealth are used interchangeably throughout the literature.

Tables

Table 9

*Examples of words in chosen LIWC categories*

| Category | Word examples |
|---|---|
| Social processes | Daughter, husband, buddy, friend, baby, adult, neighbour |
| Affective processes | Happy, cried, abandon, love, nice, sweet, worried, annoyed, crying, grief, sad |
| Cognitive processes | Cause, know, ought, effect, should, would, could, maybe, guess, always, never, stop, include, without |
| Perceptual processes | Observing, heard, feeling, listen, hearing, feel, view |
| Biological processes | Eat, blood, pain, spit, hands, clinic, flu, love, horny, dish, eat |
| Personal concerns | Job, earn, win, cook, chat, apartment, family, cash, owe, church, bury, coffin |

*Note.* Reprinted from "The Development and Psychometric Properties of LIWC2007", by Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K., 2015, p. 5. Austin, TX: University of Texas at Austin.

Table 10

*Gender, age, and consumption of the participants*

| Gender | Age | Consumption at the beginning of the treatments |
|---|---|---|
| Female | 38 | 62 |
| Female | 49 | 55 |
| Male | 49 | 50 |
| Male | 38 | 30.8 |
| Female | 27 | 32 |
| Female | 59 | 41.3 |
| Female | 37 | 50.8 |
| Male | 48 | 105.6 |
| Female | 42 | 63 |
| M | 43 | 54.5 |
| SD | 9.33 | 22.41 |

Table 11

*Exchanged emails and diary logs*

| Data set | Client | Counsellor | Diary |
|---|---|---|---|
| 1 | 7 | 21 | 15 |
| 2 | 13 | 31 | 26 |
| 3 | 17 | 33 | 141 |
| 4 | 15 | 26 | 33 |
| 5 | 13 | 35 | 25 |
| 6 | 23 | 43 | 41 |
| 7 | 12 | 33 | 27 |
| 8 | 20 | 30 | 77 |
| 9 | 19 | 32 | 16 |
| Sum | 139 | 284 | 401 |
| M | 15.44 | 31.56 | 44.56 |
| SD | 4.85 | 6.04 | 40.65 |

Table 12

*Clients' and counsellors' tokens*

| Data set | Clients' Token Average | Counsellors' Token Average |
|---|---|---|
| 1 | 1996 | 12627 |
| 2 | 3915 | 19886 |
| 3 | 4784 | 19959 |
| 4 | 2795 | 13963 |
| 5 | 1842 | 19750 |
| 6 | 3946 | 19141 |
| 7 | 4885 | 18791 |
| 8 | 4983 | 9906 |
| 9 | 10584 | 19665 |
| Sum | 39730 | 153688 |
| M | 4414.4 | 17076.4 |
| SD | 2605.34 | 3843.17 |

Table 13

*Clients' and counsellors' Cognitive, affective, and social processes*

| Data set | Clients | | | Counsellors | | |
|---|---|---|---|---|---|---|
| Processes | Cognitive (%) | Affect (%) | Social (%) | Cognitive (%) | Affect (%) | Social (%) |
| 1 | 7.57 | 5.06 | 4.26 | 10.53 | 6.63 | 10.02 |
| 2 | 11.57 | 6.85 | 5.29 | 9.94 | 6.79 | 10 |
| 3 | 8.65 | 5.48 | 4.31 | 9.58 | 6.73 | 9.99 |
| 4 | 8.62 | 4.47 | 4.54 | 9.70 | 6.35 | 9.98 |
| 5 | 5.86 | 4.02 | 2.44 | 10.70 | 6.54 | 10.25 |
| 6 | 8.34 | 4.69 | 3.52 | 9.94 | 6.50 | 10.29 |
| 7 | 9.79 | 5.71 | 6.45 | 9.71 | 6.57 | 10.22 |
| 8 | 9.75 | 5.26 | 7.79 | 8.27 | 5.69 | 11.34 |
| 9 | 9.93 | 6.09 | 7.68 | 10.23 | 6.34 | 10.14 |
| M | 8.90 | 5.29 | 5.14 | 9.84 | 6.46 | 10.25 |
| SD | 1.63 | 0.86 | 1.84 | 0.70 | 0.33 | 0.43 |

Table 14

*Case 2 – Overview of LIWC categories for case number two*

| Processes | Client (%) | Counsellor (%) |
|---|---|---|
| Affect | 6.85 | 6.79 |
| Biological | 1.63 | 1.64 |
| Cognitive | 11.57 | 9.94 |
| Perceptual | 2.5 | 1.95 |
| Personal concerns | | |
| Work | 0.97 | 1.01 |
| Leisure | 0.92 | 1.38 |
| Home | 0.18 | 0.17 |
| Money | 0.18 | 0.17 |
| Religion | Missing | 0.01 |
| Death | Missing | Missing |
| Social | 5.29 | 10 |

Table 15

*LIWC subcategories for cognitive, affective, and social processes for case number two*

| Data set | Client | | | Counsellor | | |
|---|---|---|---|---|---|---|
| Processes | Cognitive (%) | Affect (%) | Social (%) | Cognitive (%) | Affect (%) | Social (%) |
| Insight | 3.4 | | | 3.47 | | |
| Causation | 1.2 | | | 1.52 | | |
| Discrepancy | 1.66 | | | 1 | | |
| Certainty | 0.97 | | | 0.99 | | |
| Tentative | 2.58 | | | 2.33 | | |
| Positive emotions | | 5.62 | | | 5.45 | |
| Negative emotions | | 1.17 | | | 1.24 | |
| Anxiety | | 0.33 | | | 0.53 | |
| Anger | | 0.1 | | | 0.07 | |
| Sadness | | 0.33 | | | 0.24 | |
| Family | | | 0.18 | | | 0.01 |
| Friends | | | 0.31 | | | 0.11 |

Table 16

*Manual LIWC search with email snippets*

| Processes | Client's Examples | Counsellor's Examples |
|---|---|---|
| Affect | "I felt defiance, anger and total disregard. I was drinking because I was scared. I knew that that I had exceeded my own limits I set" "I can relax more easily; I'm less bored; I'm less troubled by shaking, or feeling sick; I feel less emotional pain" | "Appreciate your own efforts and the little successes." "Such negative emotions and thoughts are such a burden and would wear anyone down" |
| Cognitive | "That I am not going to be able to keep to this very rigid and structured cut down. I am thinking that my hackles are rising and that a mutinous attitude is setting in" | "Many people tend to think negative thoughts because they've learned to over the years , and they're used to it" |
| Social | "I agree with you that i need to parent myself but I found myself under immense pressure and felt myself dig my heels in and could not overcome this with any sort of rational thinking." | "You can also expect positive reactions from family, friends and acquaintances." |

# Figures

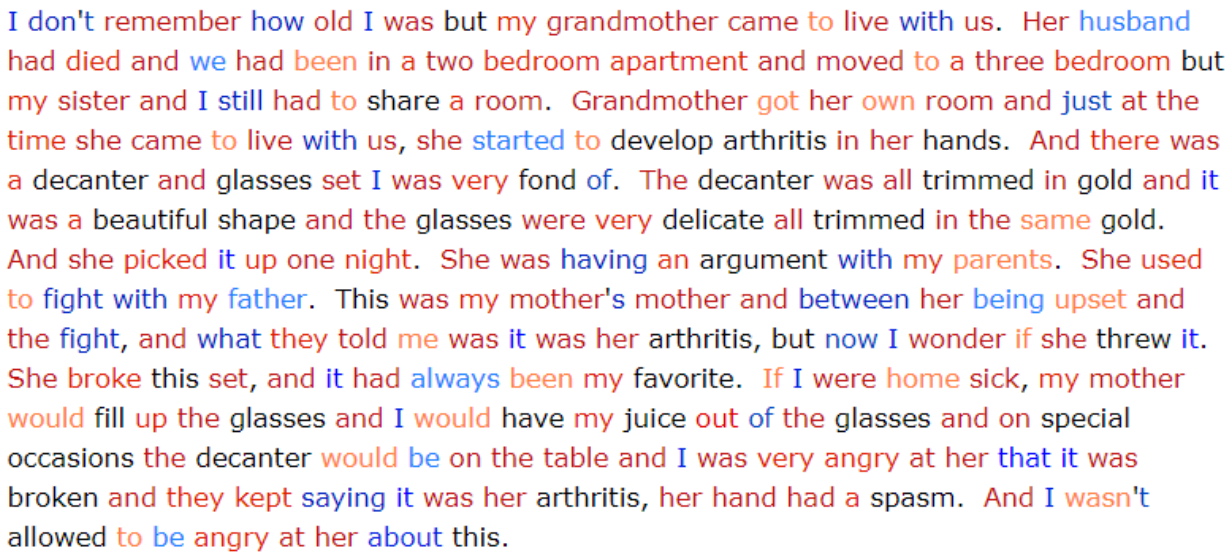



*Figure 12.* Example of RA activity as seen with WRAD dictionary

*Note.* Reprinted from The Weighted Referential Activity Dictionary (WRAD). (2019, March 6). Retrieved from The Referential Process: http://www.thereferentialprocess.org/dictionary-measures-and-computer-programs/weighted-referential-activity-dictionary-wrad

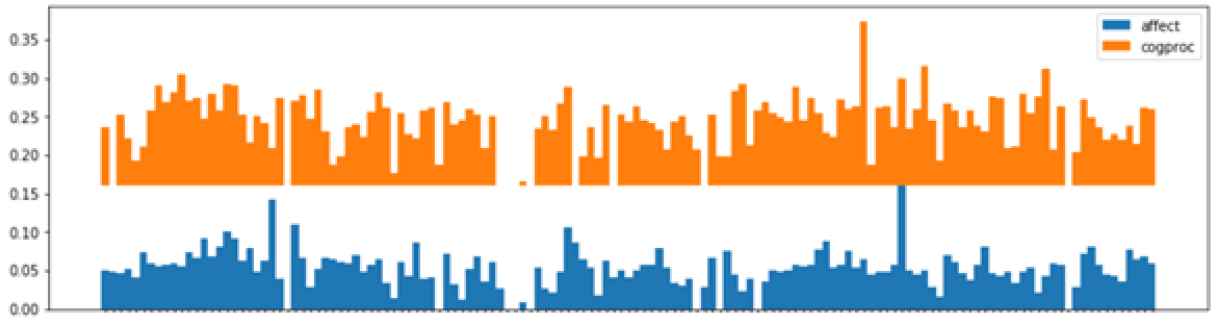



*Figure 13.* Client' cognitive and affective processes

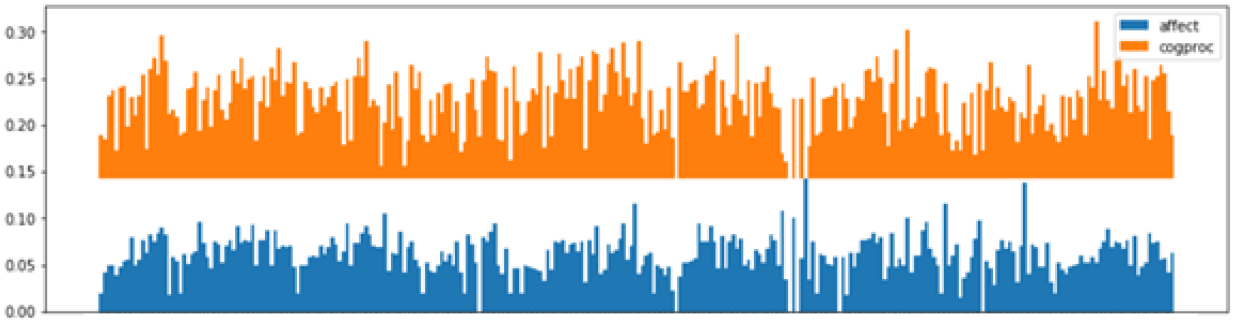



*Figure 14.* Counsellors' cognitive and affective processes

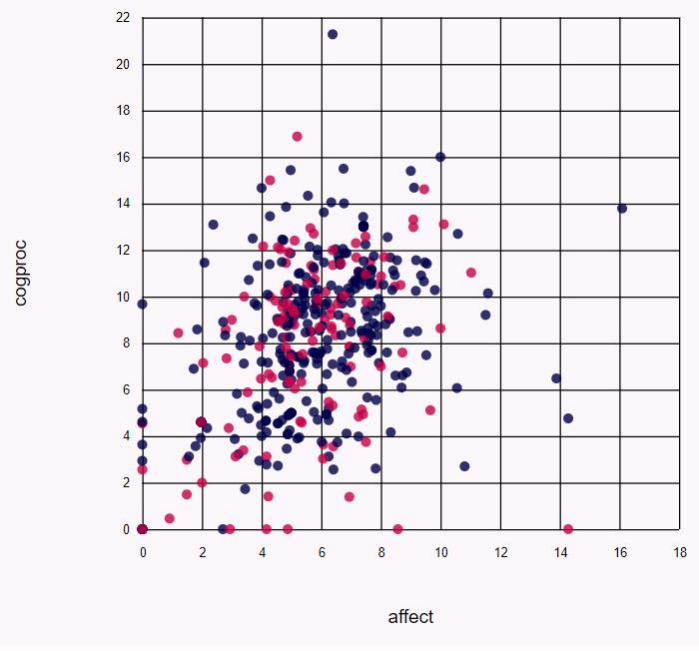



*Figure 15.* Clients' and Counsellors' tokens on scatter plot

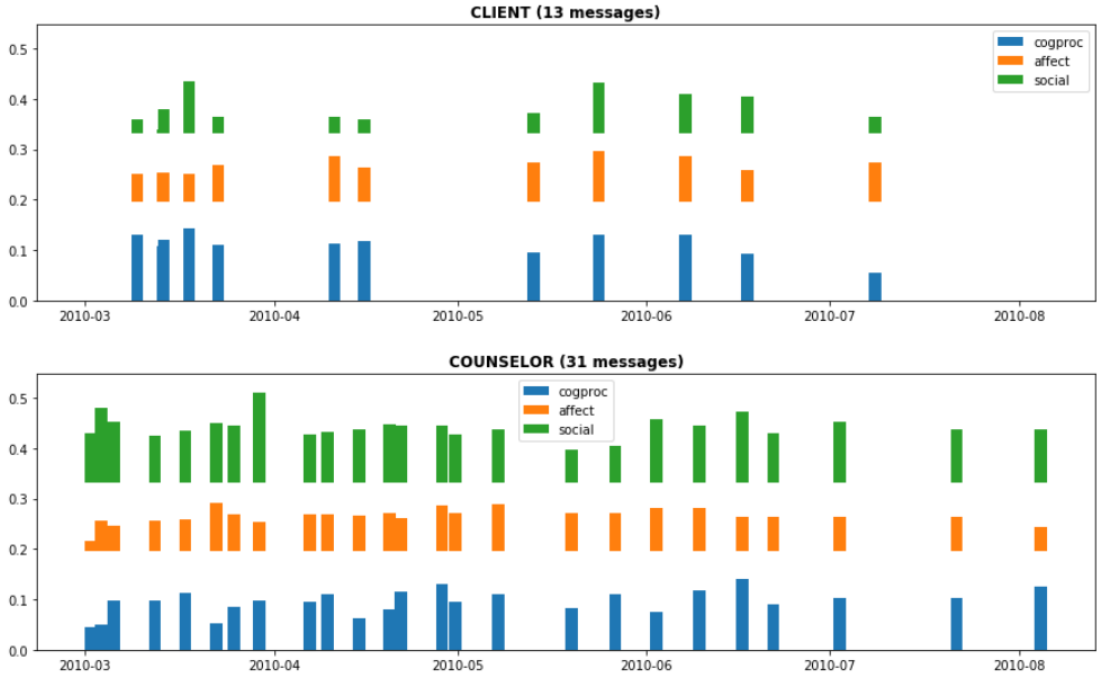



*Figure 16.* Distribution of social, cognitive, and affective processes in case number two

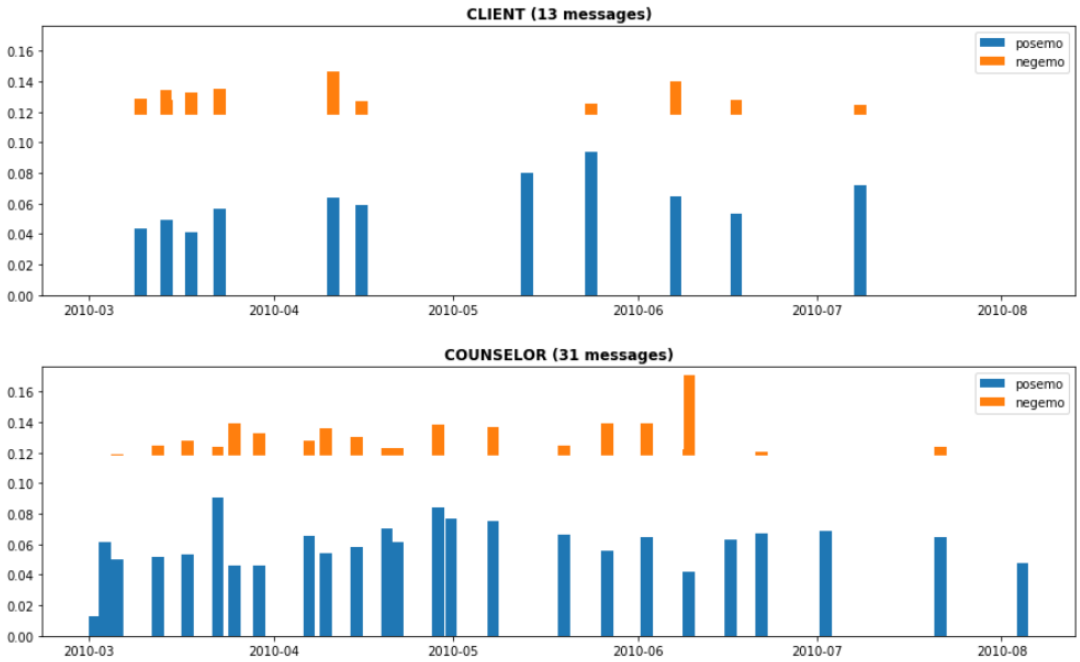



*Figure 17.* Positive and negative emotions over the time of correspondence

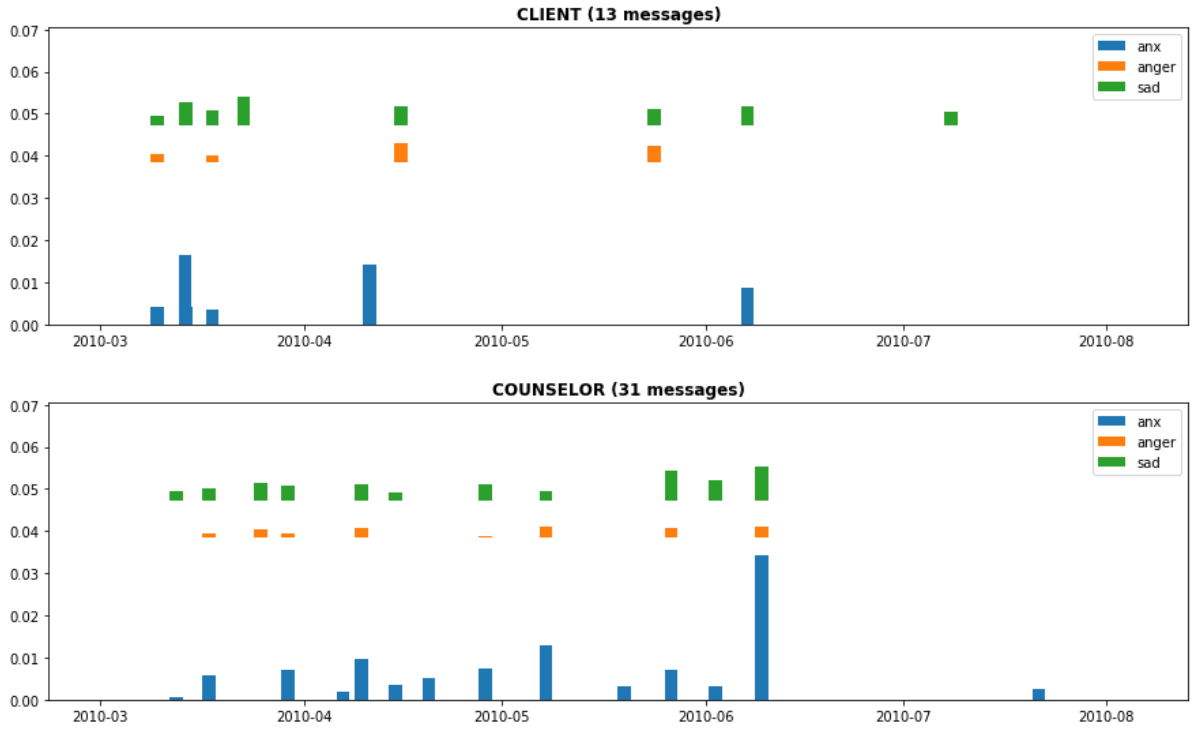



*Figure 18.* Anxiety, anger, and sadness for case two

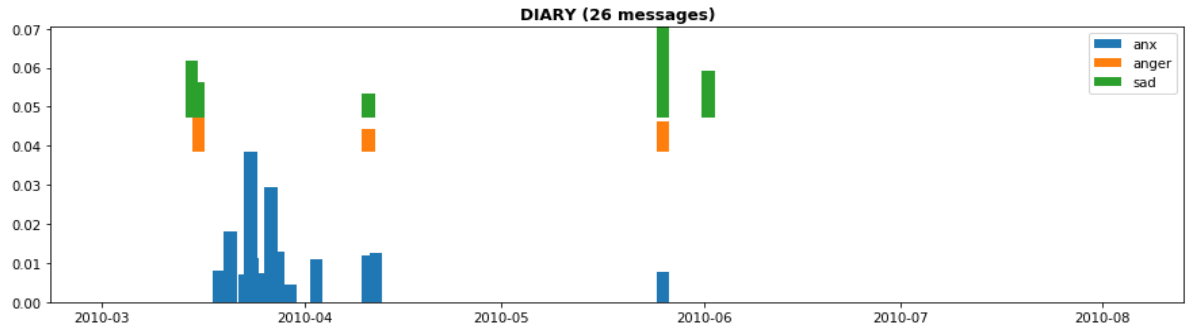



*Figure 19.* Anxiety, anger, and sadness in diaries

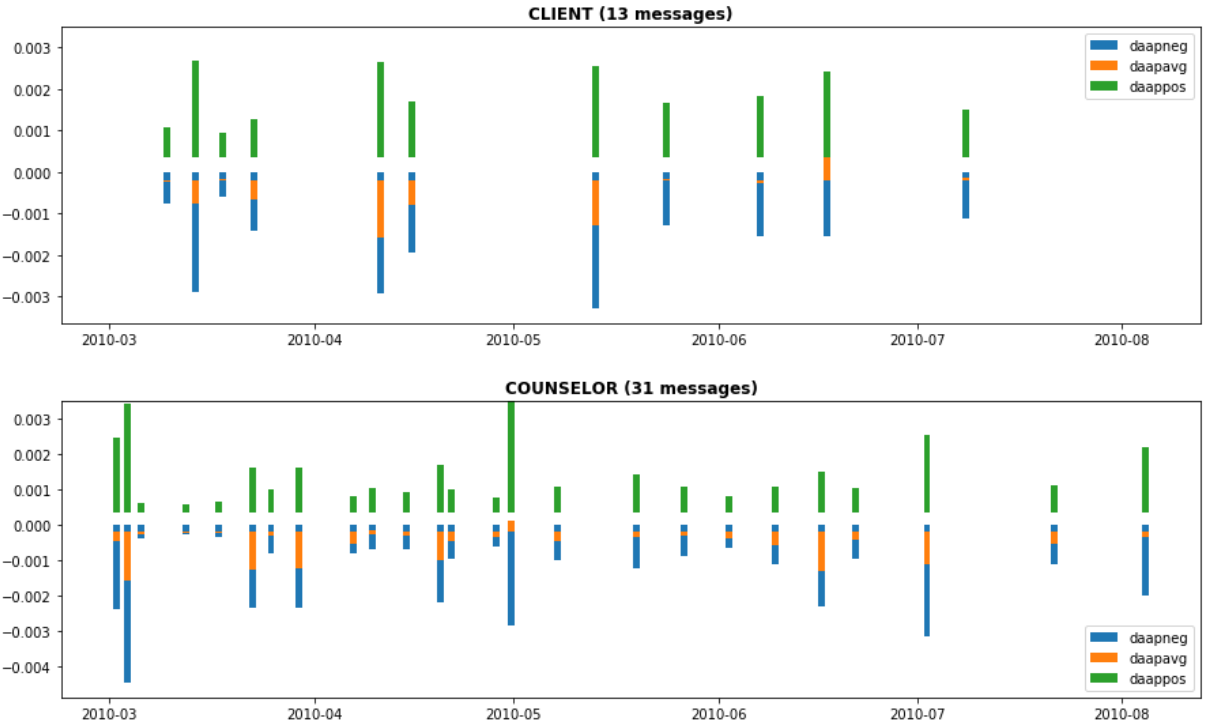



*Figure 20.* Referential Activity for case number two

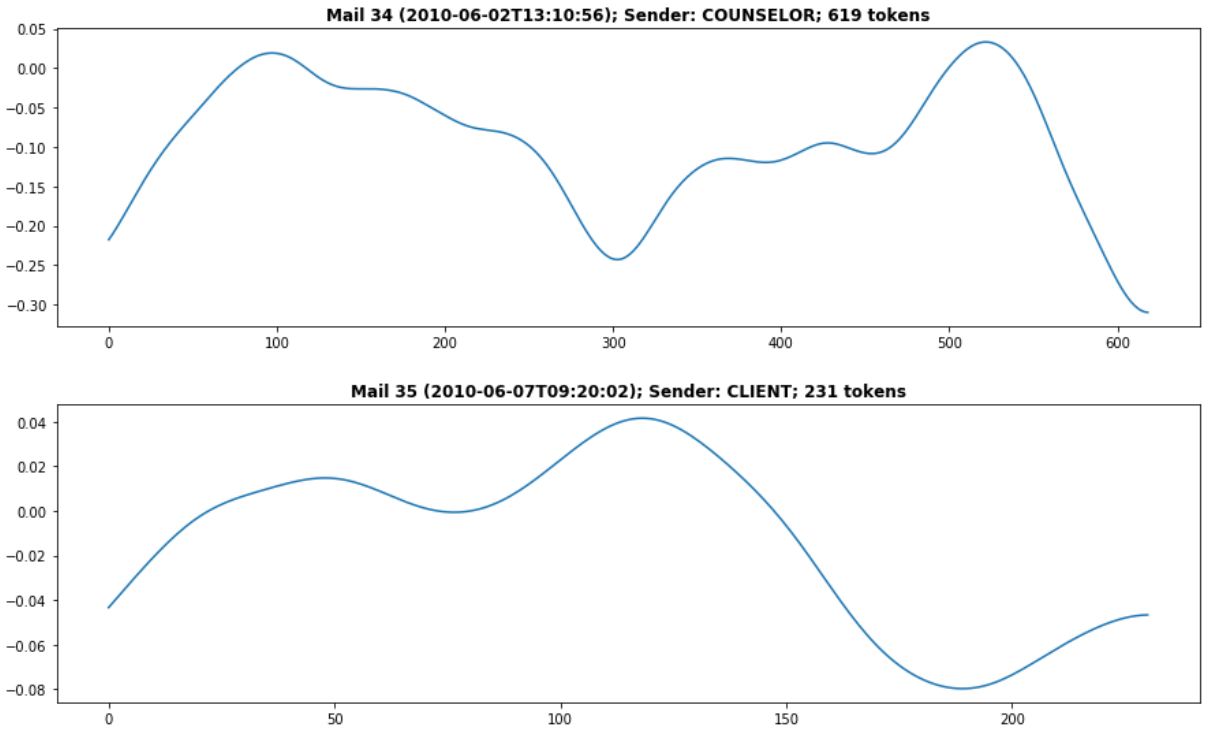



*Figure 21.* Referential Activity within given timeframe for case number two

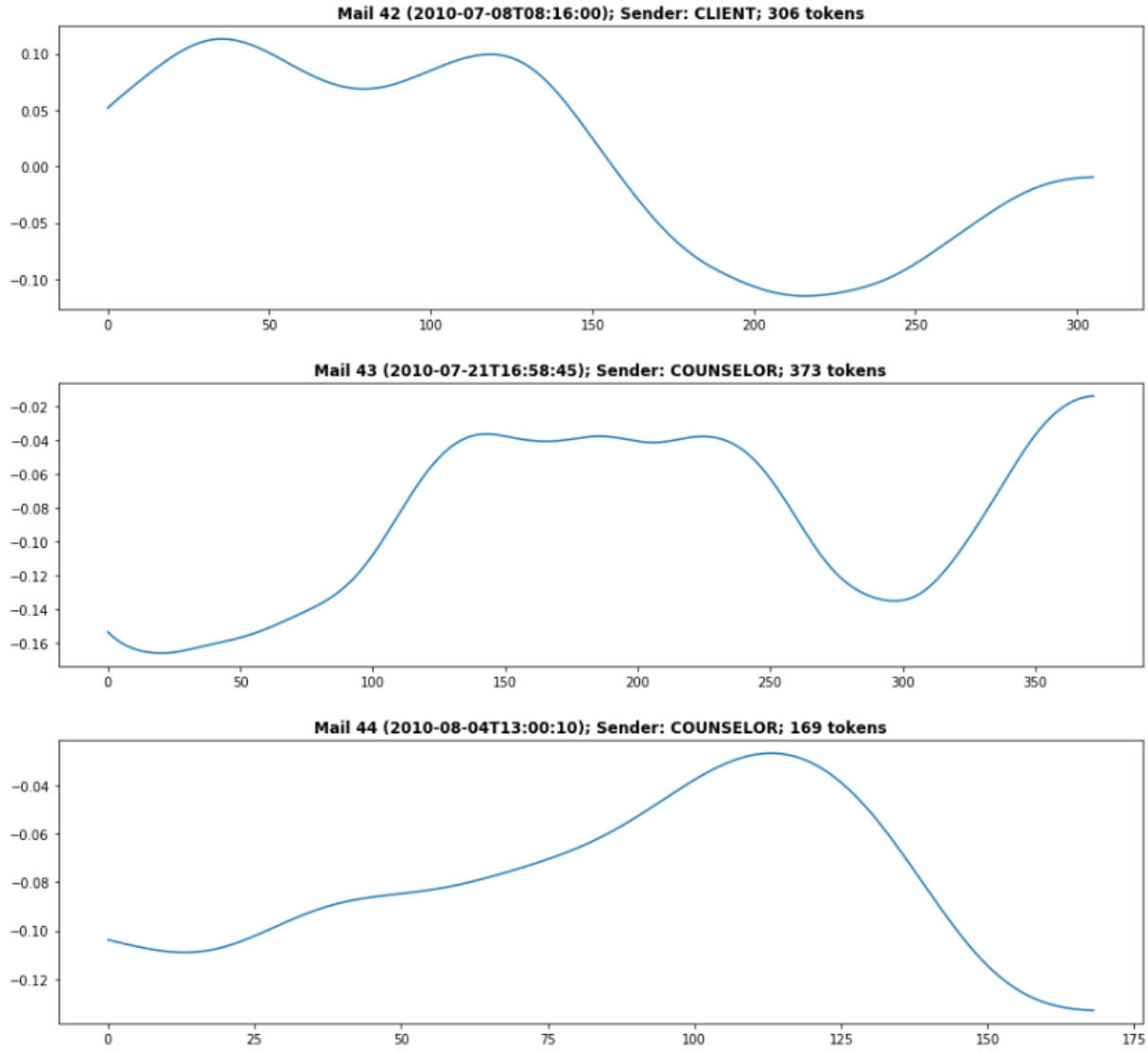
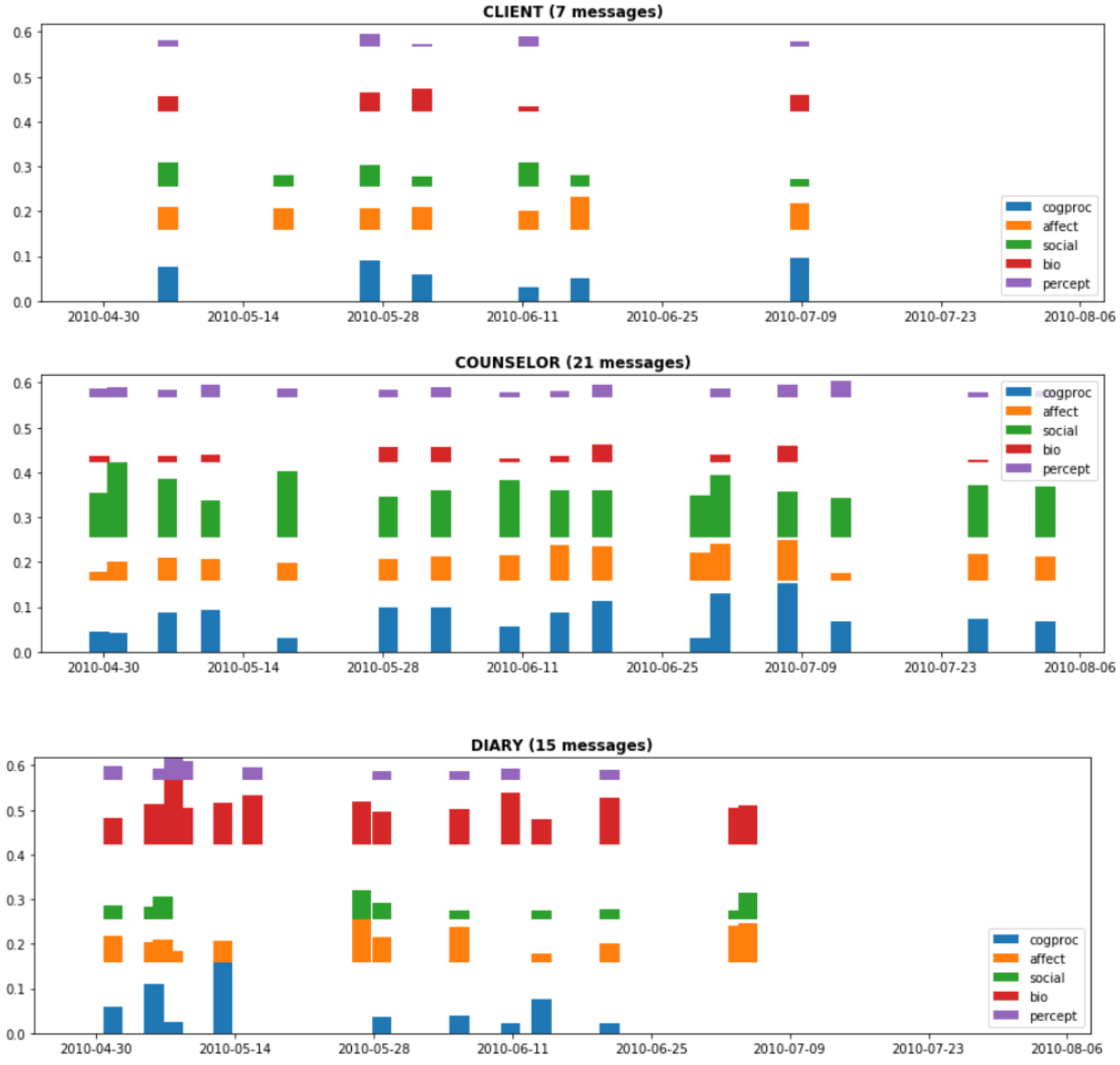



*Figure 22.* Referential Activity at the end of the treatment for client number two

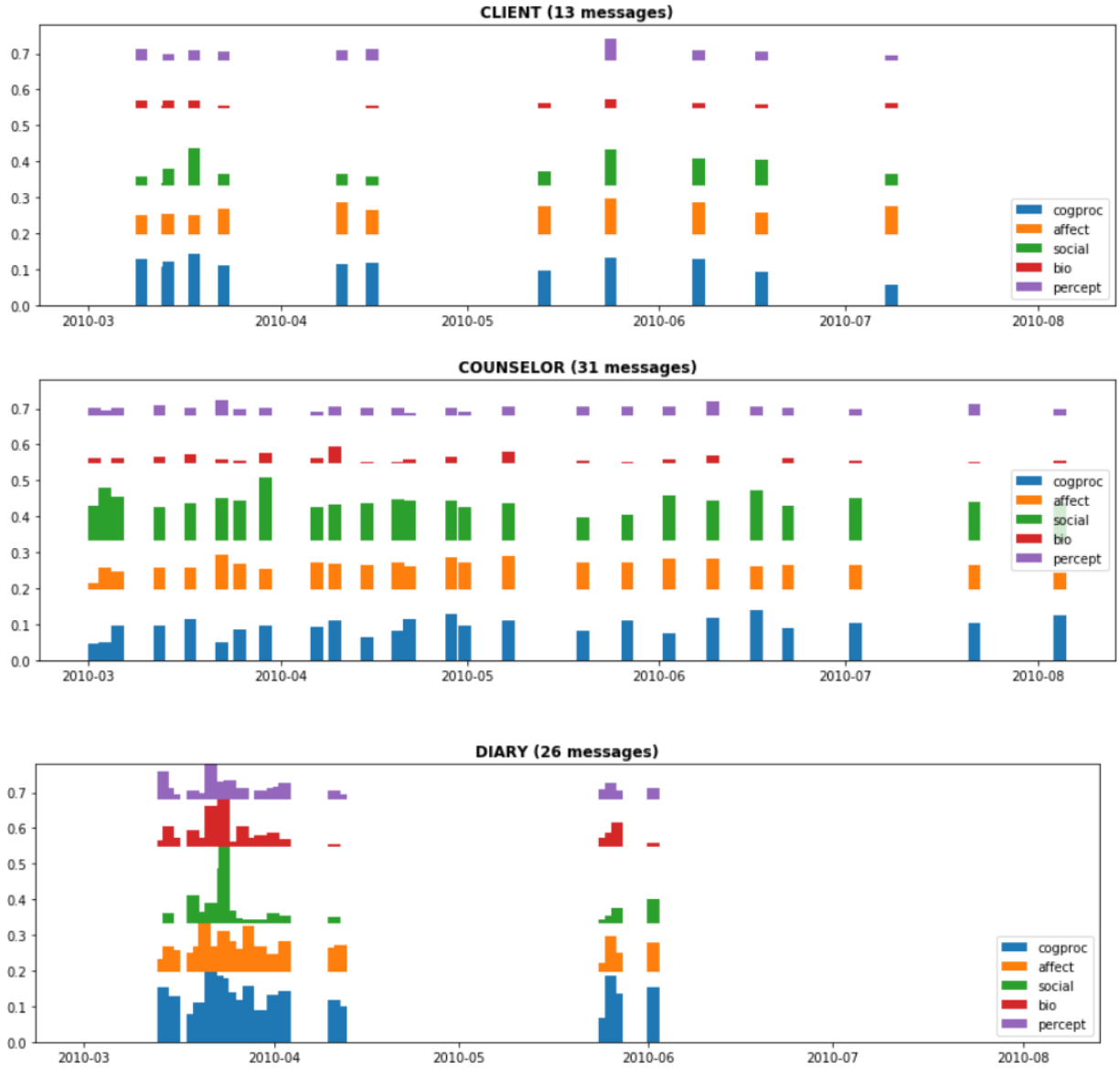
# Appendices

## Graphs of LIWC categories

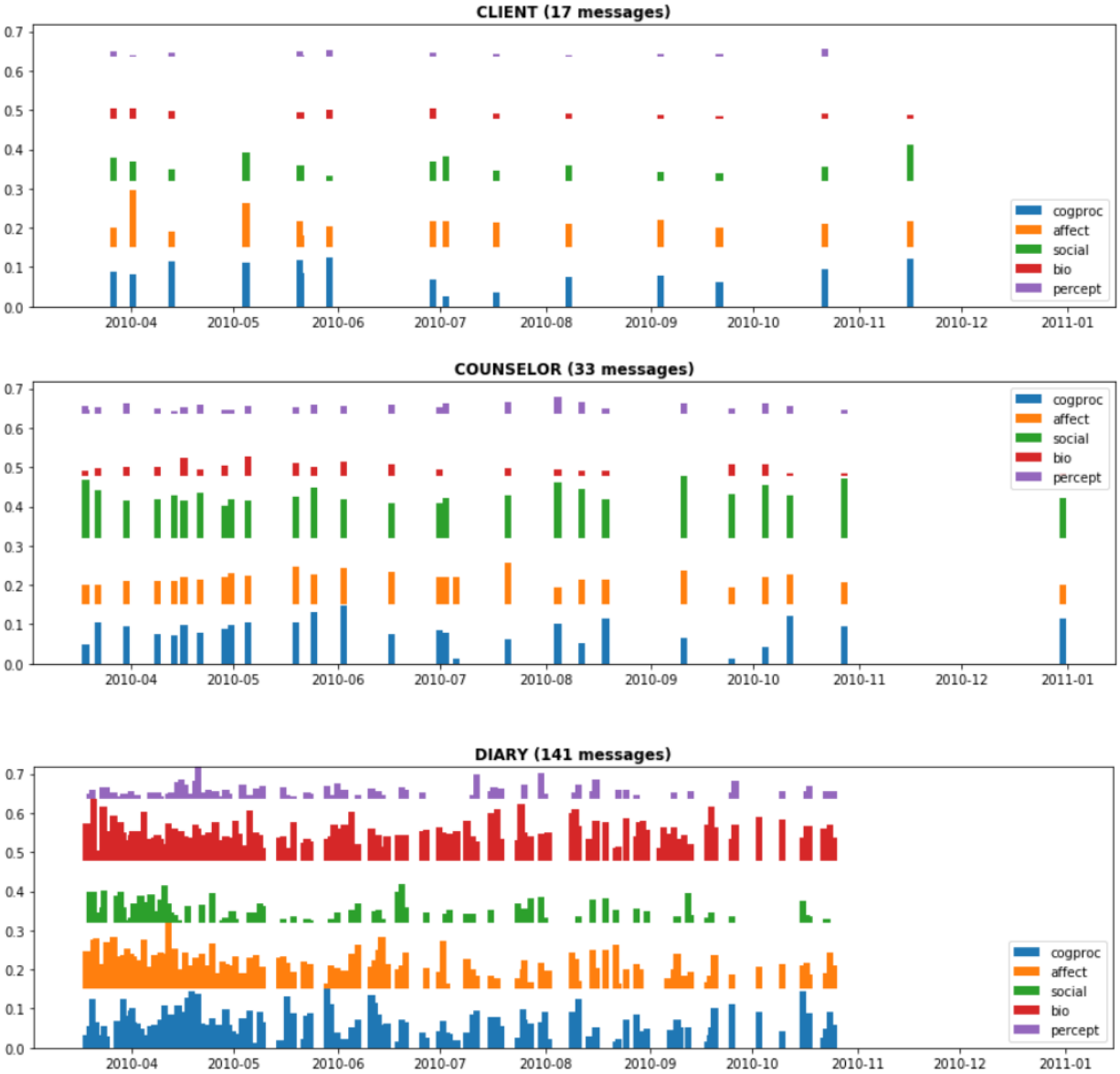
1



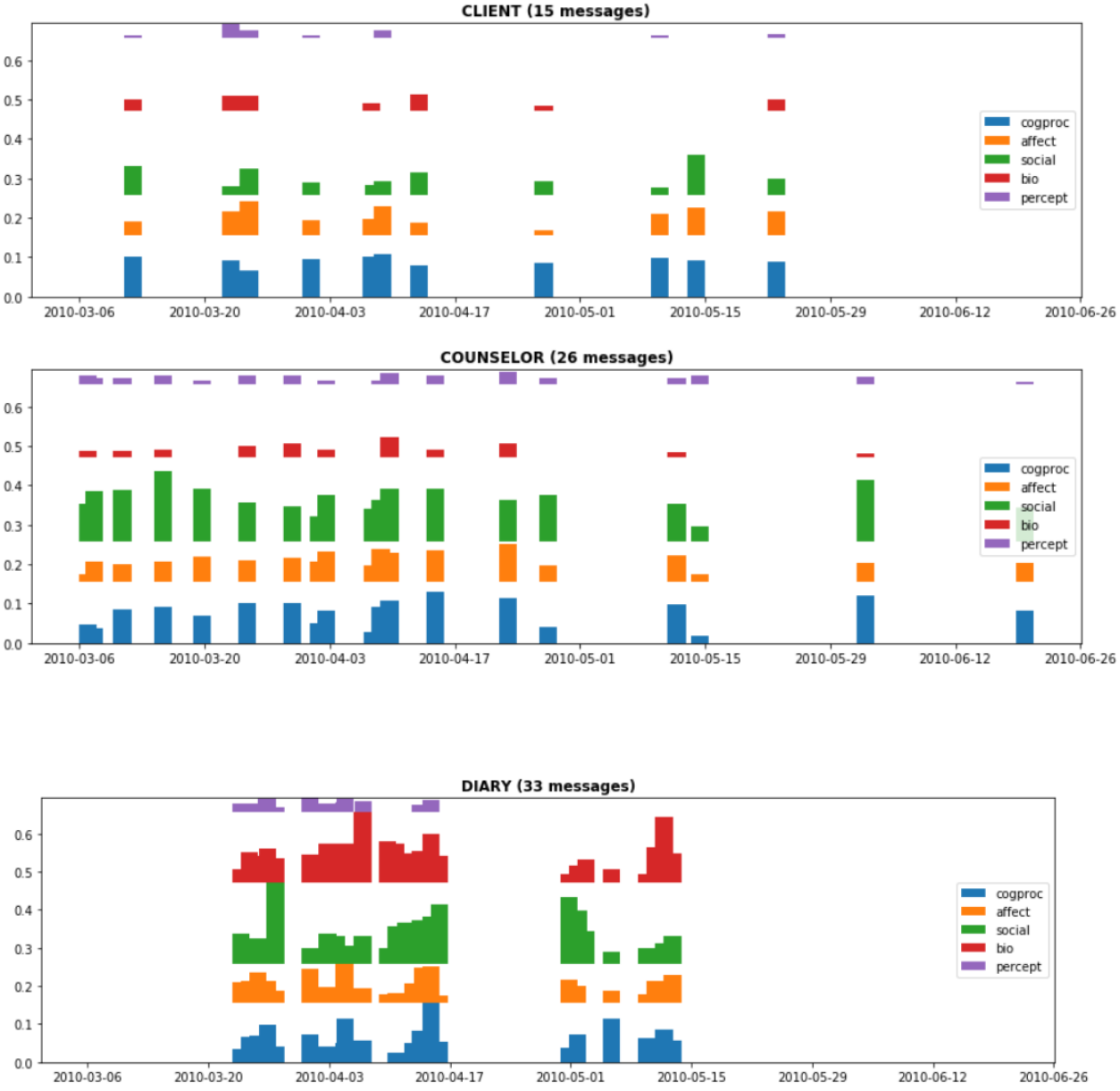
2


CLIENT (13 messages)
COUNSELOR (31 messages)
DIARY (26 messages)
cogproc
affect
social
bio
percept
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2010-03
2010-04
2010-05
2010-06
2010-07
2010-08

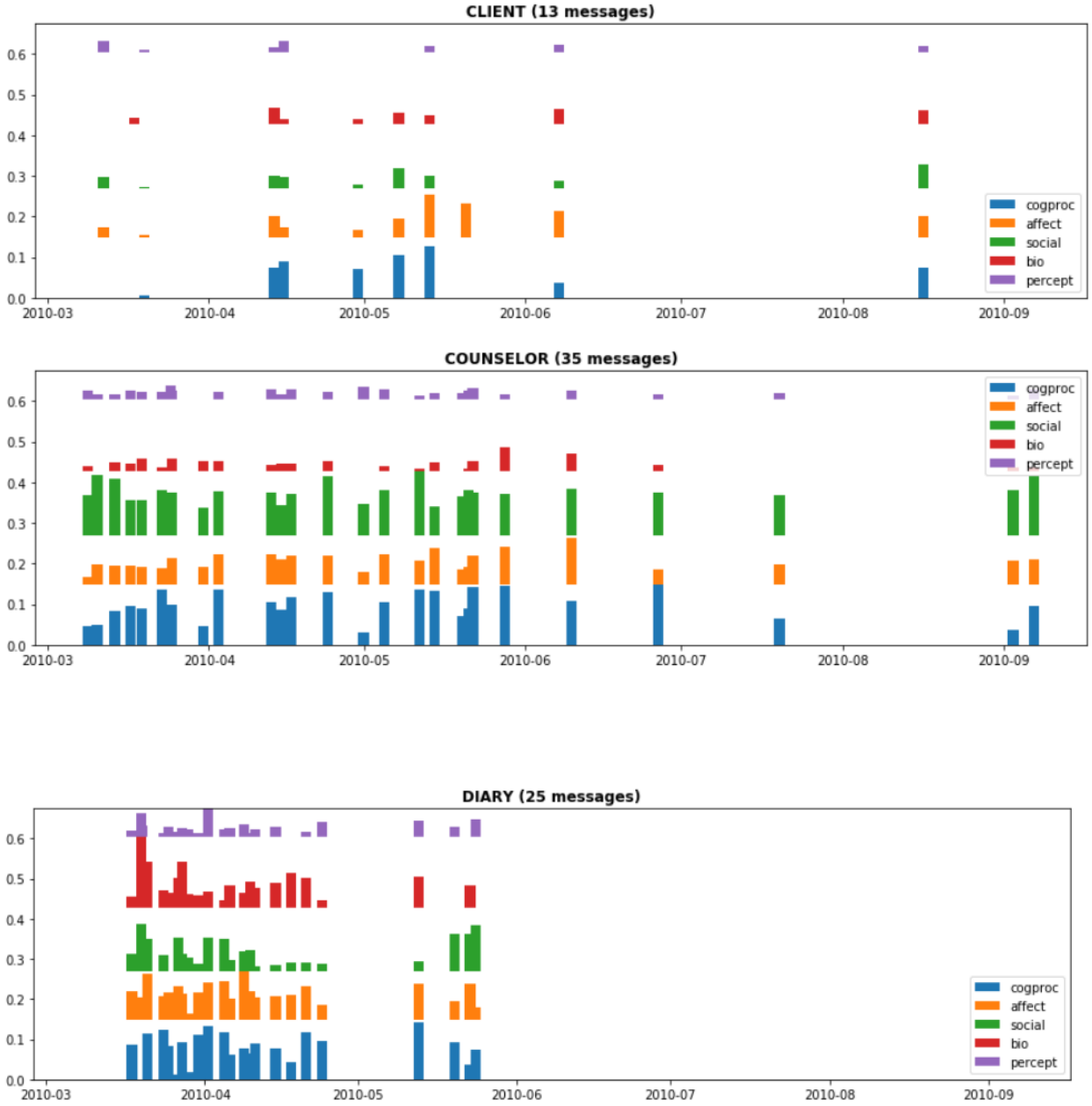
3


CLIENT (17 messages)
COUNSELOR (33 messages)
DIARY (141 messages)
cogproc
affect
social
bio
percept
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2010-04
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10
2010-11
2010-12
2011-01

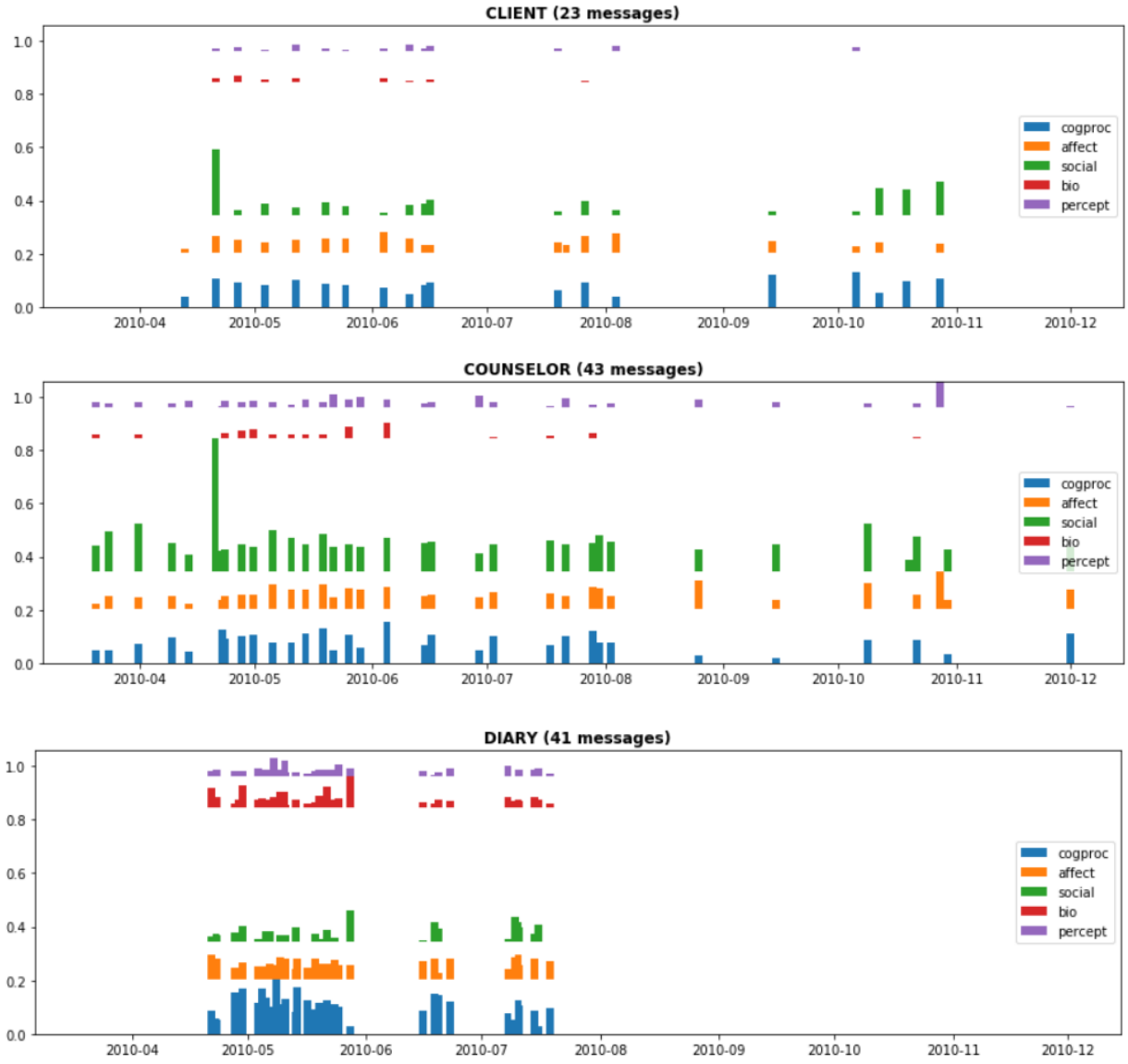
4


CLIENT (15 messages)
COUNSELOR (26 messages)
DIARY (33 messages)
cogproc
affect
social
bio
percept
0.0
0.1
0.2
0.3
0.4
0.5
0.6
2010-03-06
2010-03-20
2010-04-03
2010-04-17
2010-05-01
2010-05-15
2010-05-29
2010-06-12
2010-06-26

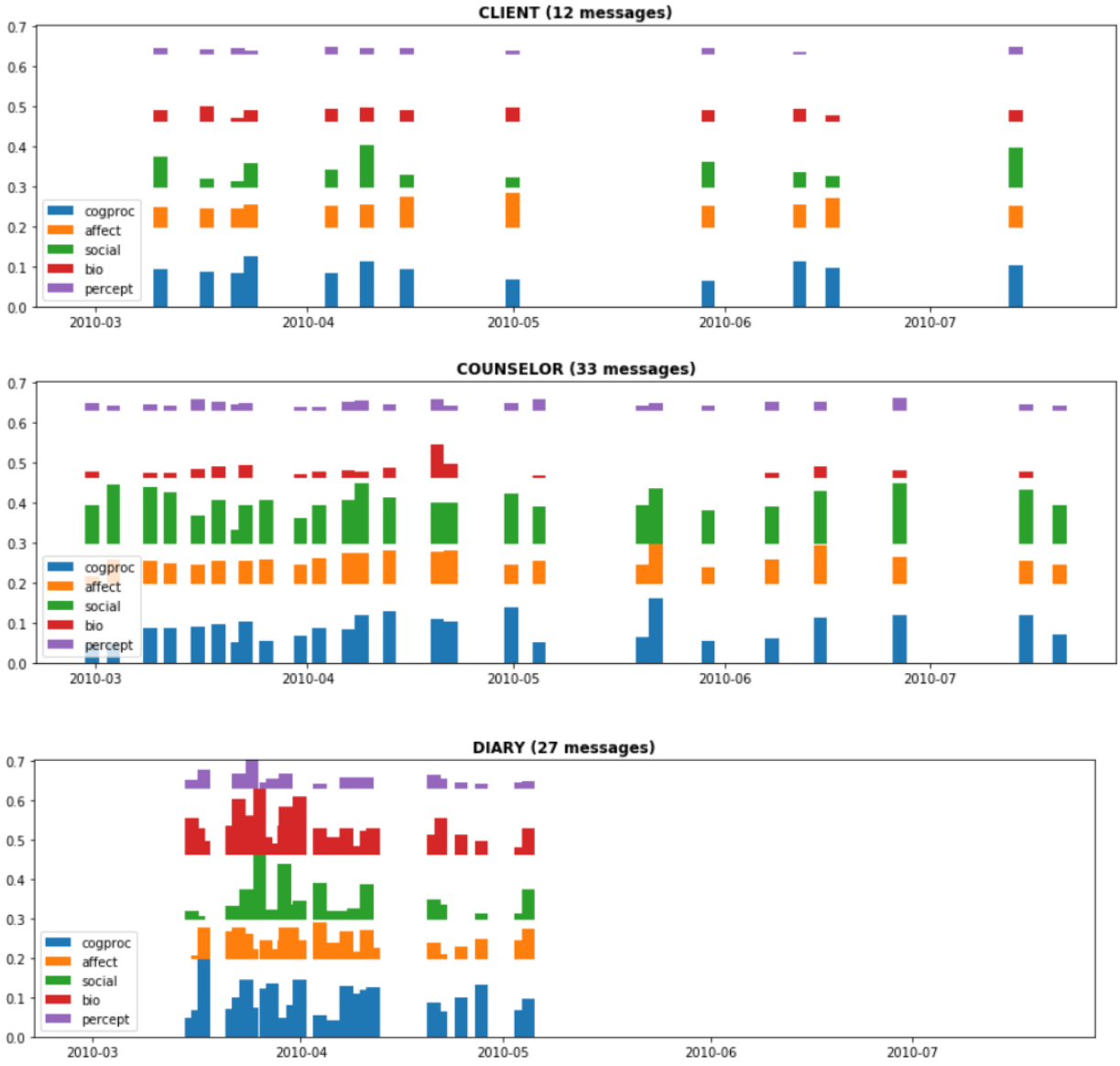
5


CLIENT (13 messages)
COUNSELOR (35 messages)
DIARY (25 messages)
cogproc
affect
social
bio
percept
0.0
0.1
0.2
0.3
0.4
0.5
0.6
2010-03
2010-04
2010-05
2010-06
2010-07
2010-08
2010-09

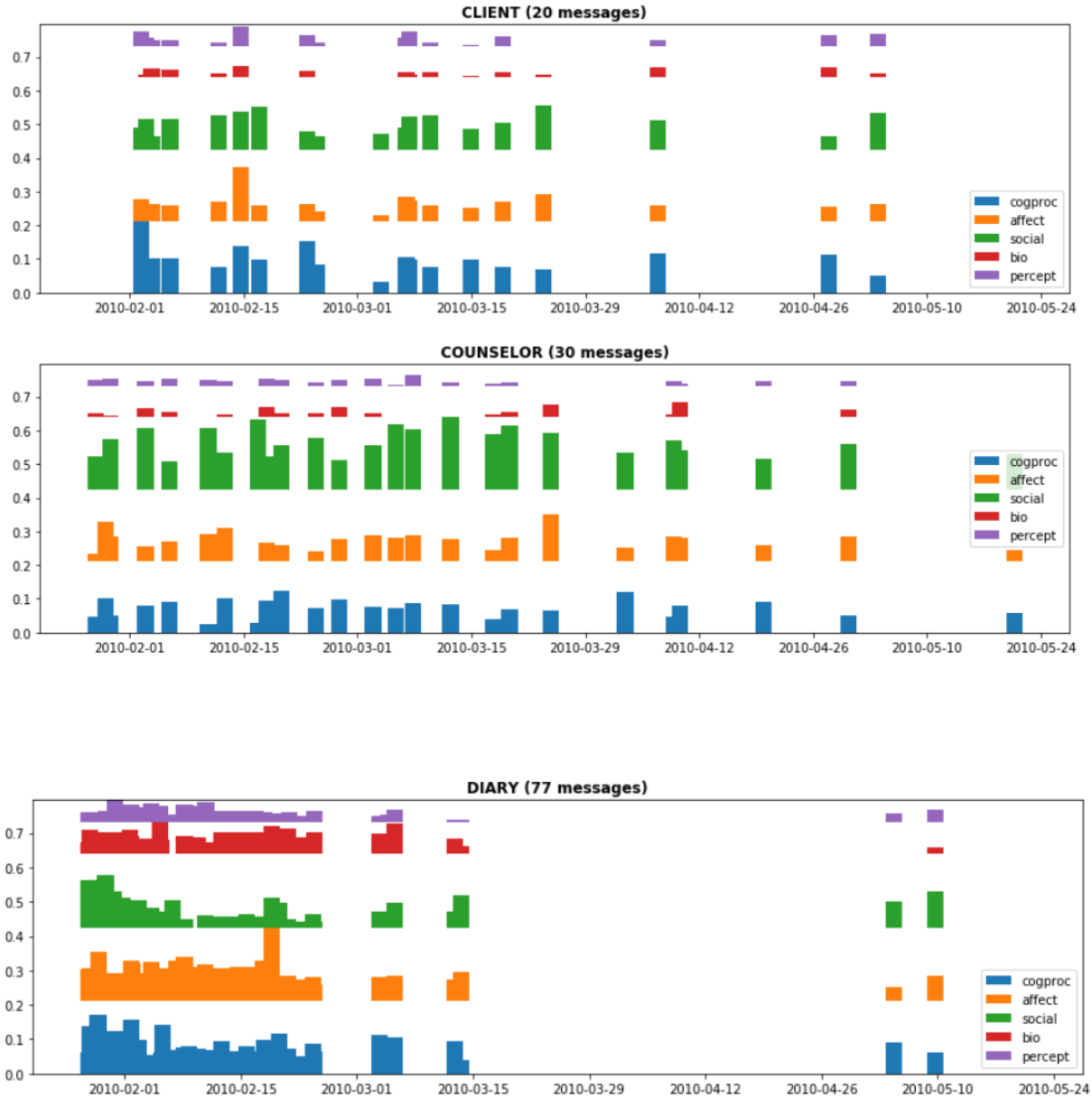
6


CLIENT (23 messages)
COUNSELOR (43 messages)
DIARY (41 messages)
cogproc
affect
social
bio
percept
0.0
0.2
0.4
0.6
0.8
1.0
2010-04
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10
2010-11
2010-12

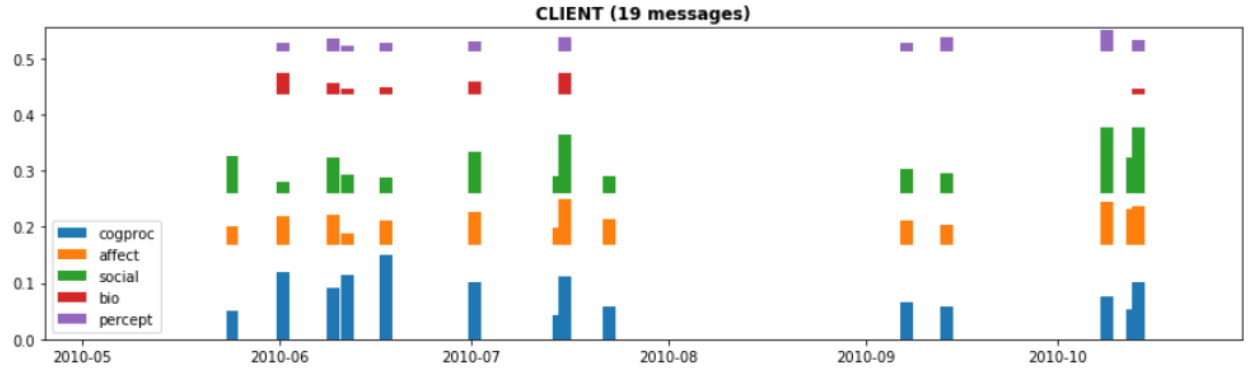
7

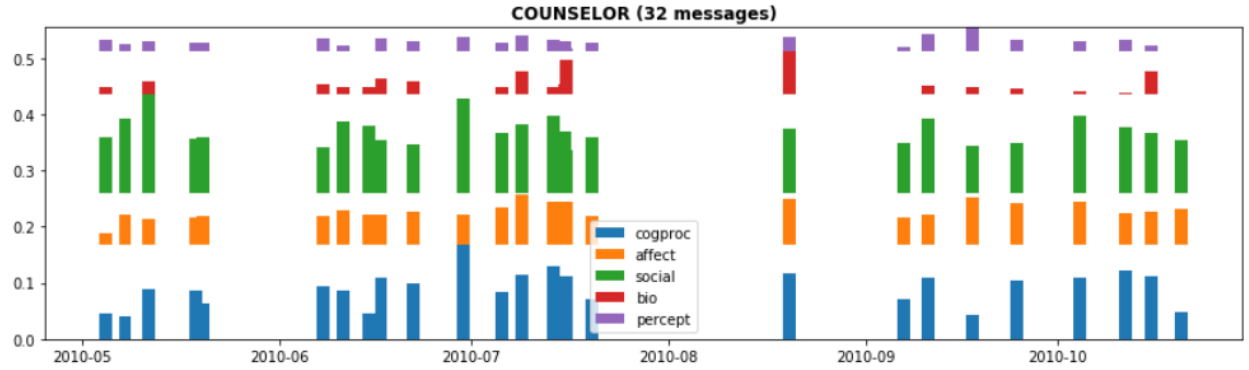

CLIENT (12 messages)
COUNSELOR (33 messages)
DIARY (27 messages)
cogproc
affect
social
bio
percept
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2010-03
2010-04
2010-05
2010-06
2010-07

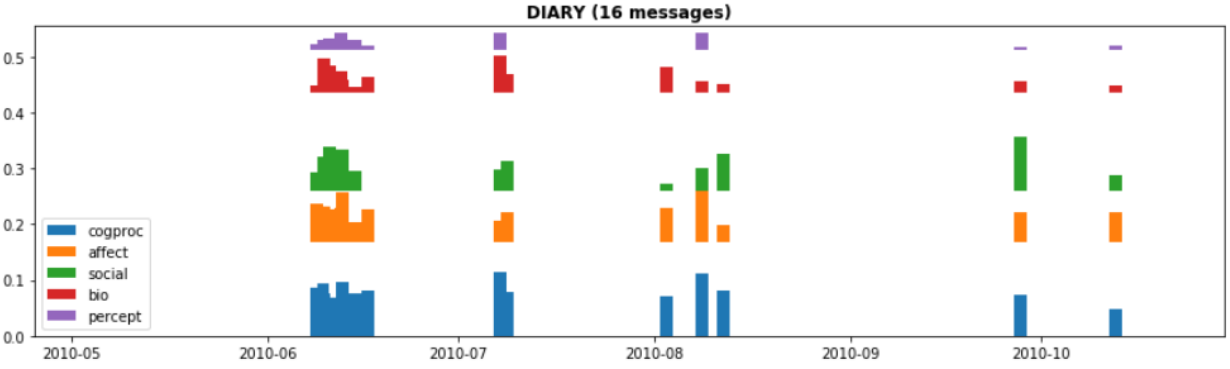
8


CLIENT (20 messages)
COUNSELOR (30 messages)
DIARY (77 messages)
cogproc
affect
social
bio
percept
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2010-02-01
2010-02-15
2010-03-01
2010-03-15
2010-03-29
2010-04-12
2010-04-26
2010-05-10
2010-05-24

9


CLIENT (19 messages)
cogproc
affect
social
bio
percept
0.0
0.1
0.2
0.3
0.4
0.5
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10


COUNSELOR (32 messages)
cogproc
affect
social
bio
percept
0.0
0.1
0.2
0.3
0.4
0.5
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10


DIARY (16 messages)
cogproc
affect
social
bio
percept
0.0
0.1
0.2
0.3
0.4
0.5
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10

LIWC Categories – Tables

| 1 | Client | Counsellor |
|---|---|---|
| Perceptual | 1.60% | 1.93% |
| Cognitive | 7.57% | 10.53% |
| Affect | 5.06% | 6.63% |
| Social | 4.26% | 10.02% |
| Biological | 3.51% | 1.83% |

| 2 | Client | Counsellor |
|---|---|---|
| Perceptual | 2.50% | 1.95% |
| Cognitive | 11.57% | 9.94% |
| Affect | 6.85% | 6.79% |
| Social | 5.29% | 10.00% |
| Biological | 1.63% | 1.64% |

| 3 | Client | Counsellor |
|---|---|---|
| Perceptual | 0.92% | 1.76% |
| Cognitive | 8.65% | 9.58% |
| Affect | 5.48% | 6.73% |
| Social | 4.31% | 9.99% |
| Biological | 1.78% | 1.97% |

| 4 | Client | Counsellor |
|---|---|---|
| Perceptual | 0.82% | 1.78% |
| Cognitive | 8.62% | 9.70% |
| Affect | 4.47% | 6.35% |
| Social | 4.54% | 9.98% |
| Biological | 2.11% | 1.76% |

| 5 | Client | Counsellor |
|---|---|---|
| Perceptual | 1.09% | 1.82% |
| Cognitive | 5.86% | 10.70% |
| Affect | 4.02% | 6.54% |
| Social | 2.44% | 10.25% |
| Biological | 1.85% | 1.73% |

| 6 | Client | Counsellor |
|---|---|---|
| Perceptual | 0.99% | 1.88% |
| Cognitive | 8.34% | 9.94% |
| Affect | 4.69% | 6.50% |
| Social | 3.52% | 10.29% |
| Biological | 0.99% | 1.69% |

| 7 | Client | Counsellor |
|---|---|---|
| Perceptual | 1.49% | 1.92% |
| Cognitive | 9.79% | 9.71% |
| Affect | 5.71% | 6.57% |
| Social | 6.45% | 10.22% |
| Biological | 2.89% | 1.71% |

| 8 | Client | Counsellor |
|---|---|---|
| Perceptual | 2.05% | 1.57% |
| Cognitive | 9.75% | 8.27% |
| Affect | 5.26% | 5.69% |
| Social | 7.79% | 11.34% |
| Biological | 1.79% | 1.98% |

| 9 | Client | Counsellor |
|---|---|---|
| Perceptual | 1.89% | 1.82% |
| Cognitive | 9.93% | 10.23% |
| Affect | 6.09% | 6.34% |
| Social | 7.68% | 10.14% |
| Biological | 1.81% | 1.82% |

LIWC Categories – Personal Concerns - Tables

1

| Personal concerns | Client | Counsellor |
|---|---|---|
| Work | 1.05% | 1.05% |
| Leisure | 2.35% | 1.63% |
| Home | 0.30% | 0.17% |
| Money | 0.45% | 0.06% |
| Religion | missing | missing |
| Death | missing | missing |

2

| Personal concerns | Client | Counsellor |
|---|---|---|
| Work | 0.97% | 1.01% |
| Leisure | 0.92% | 1.38% |
| Home | 0.18% | 0.17% |
| Money | 0.18% | 0.17% |
| Religion | missing | 0.01% |
| Death | missing | missing |

3

| Personal concerns | Client | Counsellor |
|---|---|---|
| Work | 1.30% | 1.11% |
| Leisure | 2.49% | 1.83% |
| Home | 0.40% | 0.21% |
| Money | 0.21% | 0.18% |
| Religion | missing | 0.01% |
| Death | missing | missing |

4

| Personal concerns | Client | Counsellor |
|---|---|---|
| Work | 0.68% | 1.07% |
| Leisure | 1.72% | 1.43% |
| Home | 0.54% | 0.19% |
| Money | 0.25% | 0.09% |
| Religion | missing | missing |
| Death | missing | missing |

5

| Personal concerns | Client | Counsellor |
|---|---|---|
| Work | 0.54% | 1.02% |
| Leisure | 1.14% | 1.59% |
| Home | missing | 0.11% |
| Money | missing | 0.07% |
| Religion | missing | missing |
| Death | missing | 0.01% |

6

| Personal concerns | Client | Counsellor |
|---|---|---|
| Work | 0.96% | 1.03% |
| Leisure | 0.99% | 1.32% |
| Home | 0.13% | 0.16% |
| Money | 0.15% | 0.06% |
| Religion | missing | missing |
| Death | missing | missing |

7

| Personal concerns | Client | Counsellor |
|---|---|---|
| Work | 0.84% | 1.04% |
| Leisure | 2.13% | 1.30% |
| Home | 0.51% | 0.17% |
| Money | 0.23% | 0.07% |
| Religion | missing | missing |
| Death | 0.02% | missing |

8

| Personal concerns | Client | Counsellor |
|---|---|---|
| Work | 0.90% | 1.05% |
| Leisure | 0.46% | 1.61% |
| Home | 0.34% | 0.14% |
| Money | 0.50% | 0.08% |
| Religion | missing | missing |
| Death | 0.02% | missing |

9

| Personal concerns | Client | Counsellor |
|---|---|---|
| Work | 0.95% | 0.88% |
| Leisure | 1.63% | 1.59% |
| Home | 0.22% | 0.19% |
| Money | 0.09% | 0.11% |
| Religion | missing | 0.01% |
| Death | 0.01% | 0.01% |

LIWC Categories – Cognitive processes – Tables

1

| Cognitive | Client | Counsellor |
|---|---|---|
| Insight | 1.40% | 3.60% |
| Causation | 0.75% | 1.59% |
| Discrepancy | 1.25% | 0.79% |
| Certainty | 0.50% | 1.08% |
| Tentative | 1.75% | 2.58% |

2

| Cognitive | Client | Counsellor |
|---|---|---|
| Insight | 3.4% | 3.47% |
| Causation | 1.2% | 1.52% |
| Discrepancy | 1.66% | 1% |
| Certainty | 0.97% | 0.99% |
| Tentative | 2.58% | 2.33% |

3

| Cognitive | Client | Counsellor |
|---|---|---|
| Insight | 1.51% | 3.03% |
| Causation | 0.96% | 1.47% |
| Discrepancy | 1.21% | 0.96% |
| Certainty | 0.98% | 1.01% |
| Tentative | 2.47% | 2.19% |

4

| Cognitive | Client | Counsellor |
|---|---|---|
| Insight | 1.79% | 3.36% |
| Causation | 1.18% | 1.50% |
| Discrepancy | 0.93% | 0.79% |
| Certainty | 0.89% | 0.96% |
| Tentative | 1.68% | 2.25% |

5

| Cognitive | Client | Counsellor |
|---|---|---|
| Insight | 1.30% | 3.42% |
| Causation | 0.54% | 1.43% |
| Discrepancy | 0.98% | 1.02% |
| Certainty | 0.33% | 1.01% |
| Tentative | 1.14% | 2.77% |

6

| Cognitive | Client | Counsellor |
|---|---|---|
| Insight | 1.98% | 3.13% |
| Causation | 1.22% | 1.49% |
| Discrepancy | 0.96% | 0.86% |
| Certainty | 1.04% | 0.95% |
| Tentative | 1.90% | 2.59% |

7

| Cognitive | Client | Counsellor |
|---|---|---|
| Insight | 2.29% | 3.18% |
| Causation | 1.21% | 1.46% |
| Discrepancy | 1.19% | 0.90% |
| Certainty | 1.11% | 1.01% |
| Tentative | 2.52% | 2.39% |

8

| Cognitive | Client | Counsellor |
|---|---|---|
| Insight | 2.53% | 2.70% |
| Causation | 0.80% | 1.30% |
| Discrepancy | 1.73% | 1.01% |
| Certainty | 1.04% | 0.54% |
| Tentative | 2.09% | 2.54% |

9

| Cognitive | Client | Counsellor |
|---|---|---|
| Insight | 2.65% | 3.32% |
| Causation | 1.13% | 1.41% |
| Discrepancy | 1.11% | 0.95% |
| Certainty | 0.99% | 0.96% |
| Tentative | 2.59% | 2.60% |

LIWC Categories – Affective processes – Tables

1

| Affective | Client | Counsellor |
|---|---|---|
| Positive emotion | 4.06% | 5.53% |
| Negative emotion | 1.00% | 1.06% |
| Anxiety | 0.25% | 0.48% |
| Anger | missing | 0.05% |
| Sadness | 0.20% | 0.15% |

2

| Affective | Client | Counsellor |
|---|---|---|
| Positive emotion | 5.62% | 5.45% |
| Negative emotion | 1.17% | 1.24% |
| Anxiety | 0.33% | 0.53% |
| Anger | 0.10% | 0.07% |
| Sadness | 0.33% | 0.24% |

3

| Affective | Client | Counsellor |
|---|---|---|
| Positive emotion | 4.85% | 5.77% |
| Negative emotion | 0.61% | 0.92% |
| Anxiety | 0.06% | 0.39% |
| Anger | 0.04% | 0.08% |
| Sadness | 0.19% | 0.14% |

4

| Affective | Client | Counsellor |
|---|---|---|
| Positive emotion | 3.76% | 5.26% |
| Negative emotion | 0.57% | 1.06% |
| Anxiety | 0.11% | 0.51% |
| Anger | 0.04% | 0.06% |
| Sadness | 0.14% | 0.16% |

5

| Affective | Client | Counsellor |
|---|---|---|
| Positive emotion | 3.31% | 5.59% |
| Negative emotion | 0.65% | 0.88% |
| Anxiety | 0.33% | 0.39% |
| Anger | missing | 0.06% |
| Sadness | missing | 0.10% |

6

| Affective | Client | Counsellor |
|---|---|---|
| Positive emotion | 4.08% | 5.50% |
| Negative emotion | 0.56% | 0.96% |
| Anxiety | 0.05% | 0.31% |
| Anger | missing | 0.07% |
| Sadness | 0.15% | 0.14% |

7

| Affective | Client | Counsellor |
|---|---|---|
| Positive emotion | 4.65% | 5.25% |
| Negative emotion | 1.02% | 1.23% |
| Anxiety | 0.20% | 0.58% |
| Anger | 0.14% | 0.07% |
| Sadness | 0.27% | 0.16% |

8

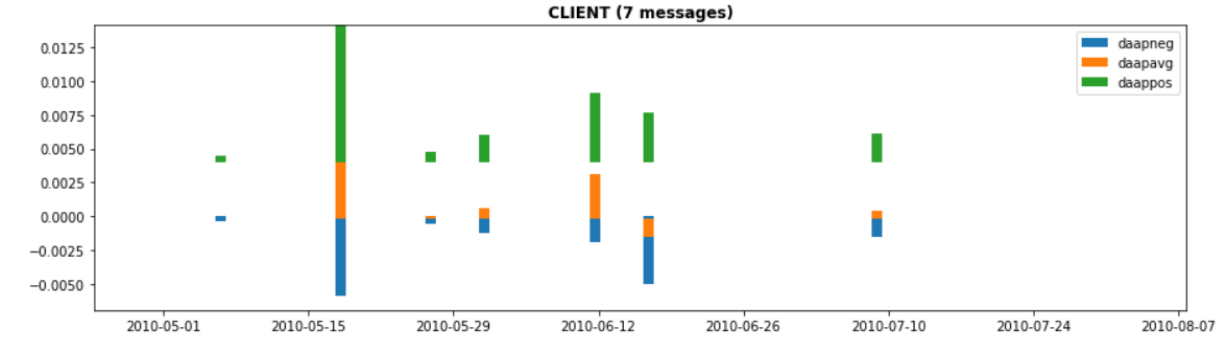
| Affective | Client | Counsellor |
|---|---|---|
| Positive emotion | 3.99% | 4.93% |
| Negative emotion | 1.20% | 0.73% |
| Anxiety | 0.24% | 0.23% |
| Anger | 0.14% | 0.03% |
| Sadness | 0.22% | 0.06% |

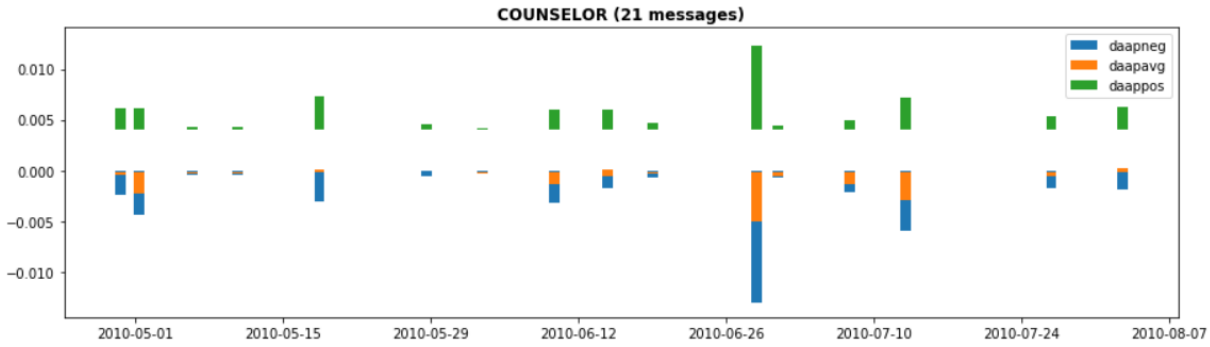
9

| Affective | Client | Counsellor |
|---|---|---|
| Positive emotion | 4.86% | 5.20% |
| Negative emotion | 1.16% | 1.08% |
| Anxiety | 0.43% | 0.42% |
| Anger | 0.09% | 0.07% |
| Sadness | 0.17% | 0.14% |

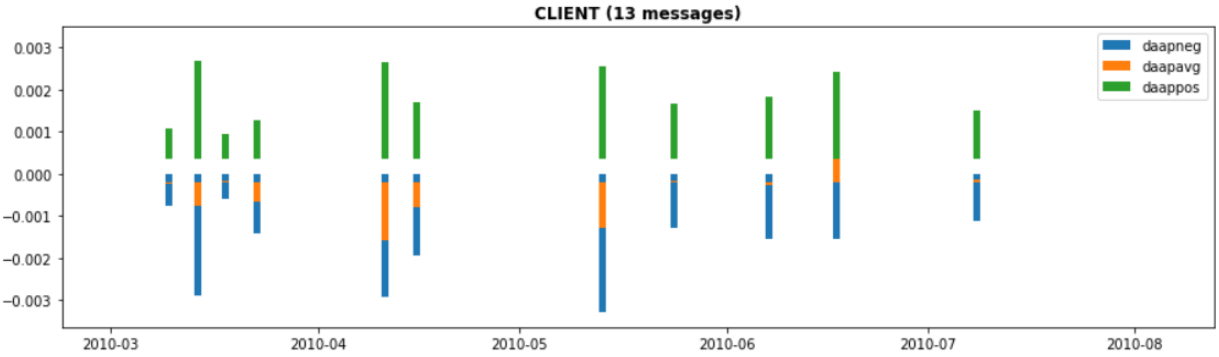
DAAP Measures

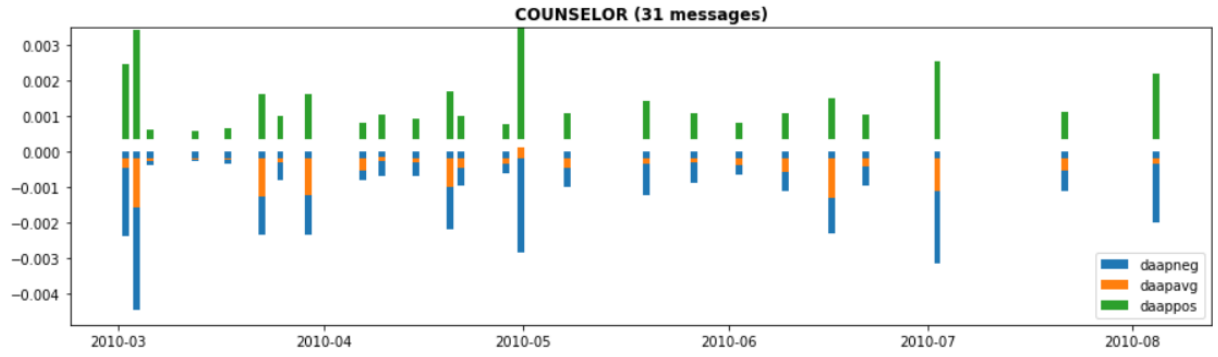
1



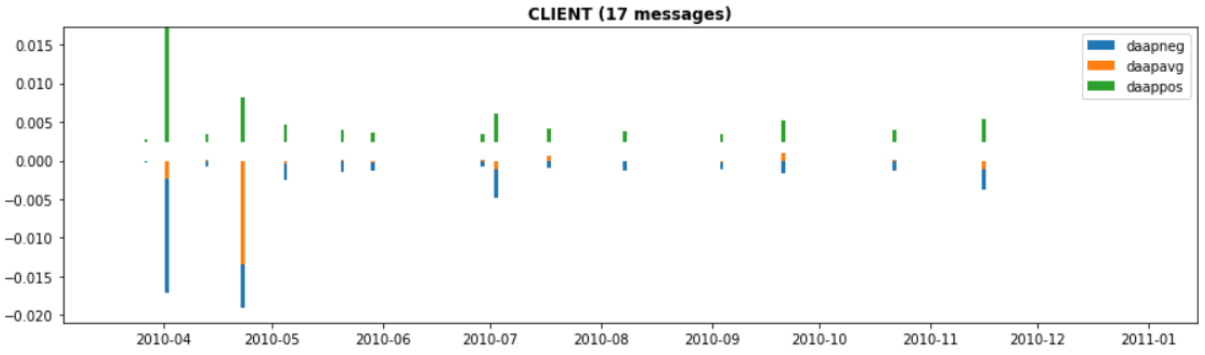



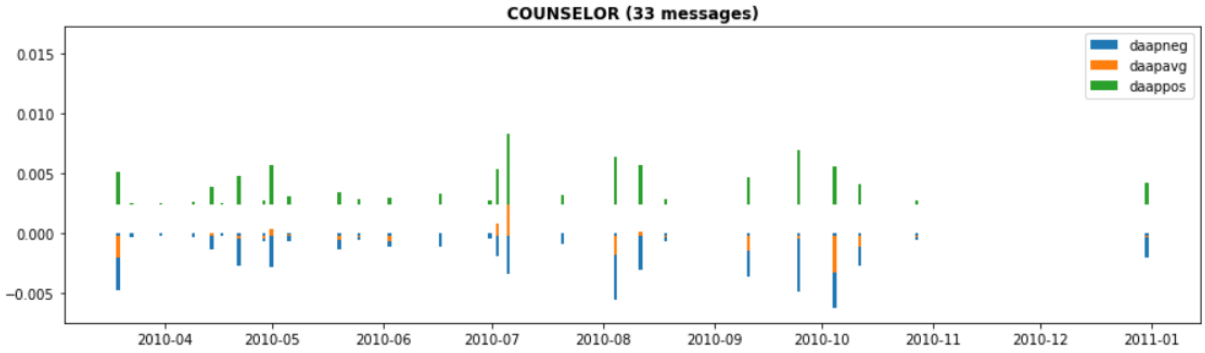
2


CLIENT (13 messages)
0.003
0.002
0.001
0.000
−0.001
−0.002
−0.003
2010-03
2010-04
2010-05
2010-06
2010-07
2010-08
daapneg
daapavg
daappos

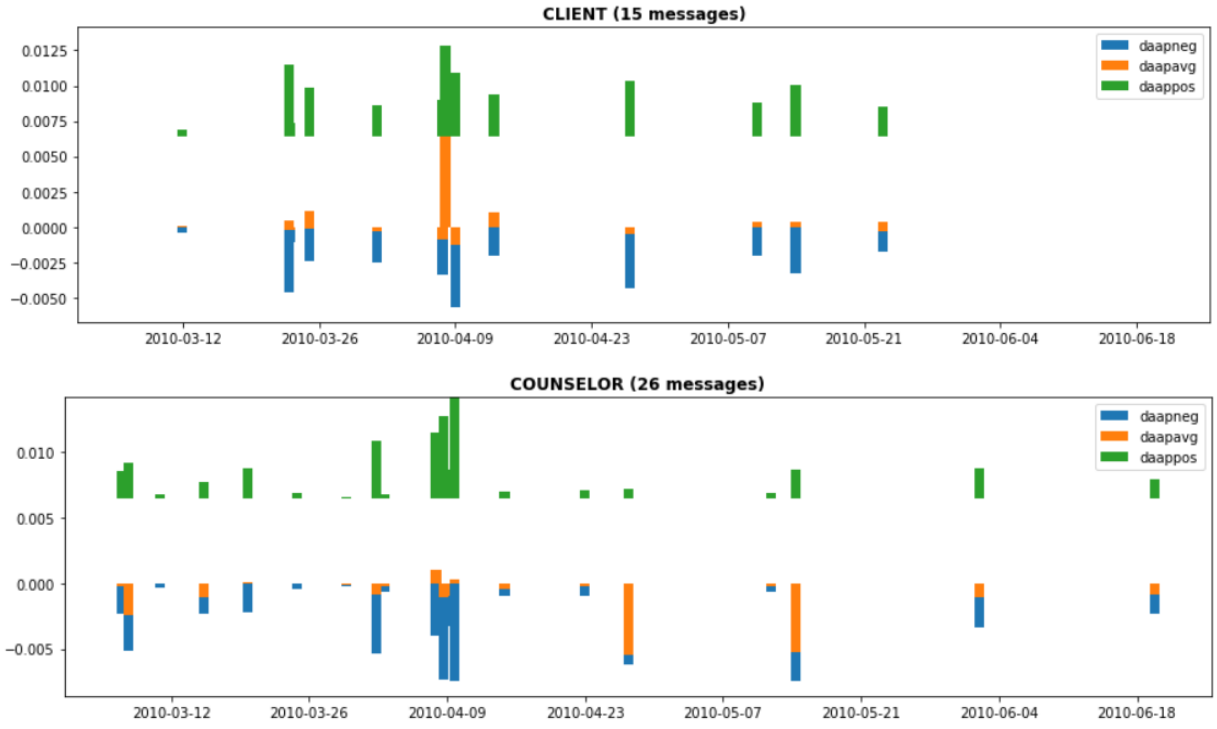

COUNSELOR (31 messages)
0.003
0.002
0.001
0.000
−0.001
−0.002
−0.003
−0.004
2010-03
2010-04
2010-05
2010-06
2010-07
2010-08
daapneg
daapavg
daappos

3

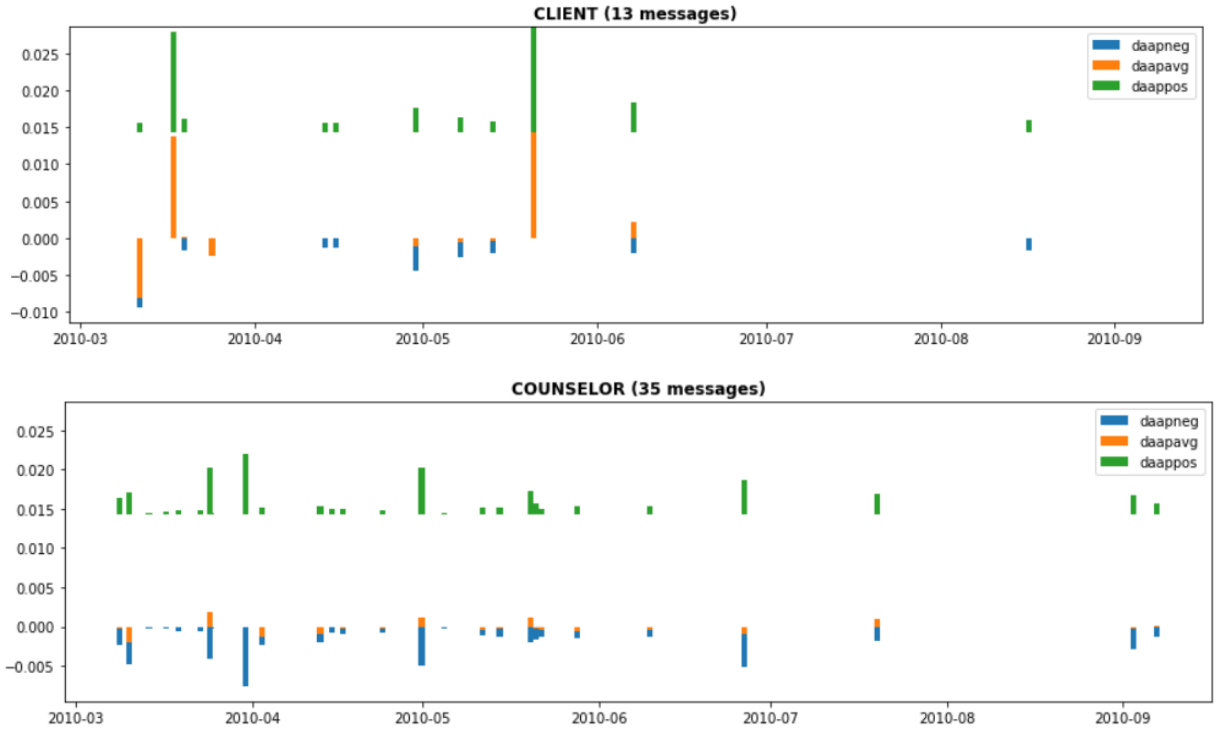

CLIENT (17 messages)
daapneg
daapavg
daappos
0.015
0.010
0.005
0.000
−0.005
−0.010
−0.015
−0.020
2010-04
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10
2010-11
2010-12
2011-01


COUNSELOR (33 messages)
daapneg
daapavg
daappos
0.015
0.010
0.005
0.000
−0.005
2010-04
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10
2010-11
2010-12
2011-01

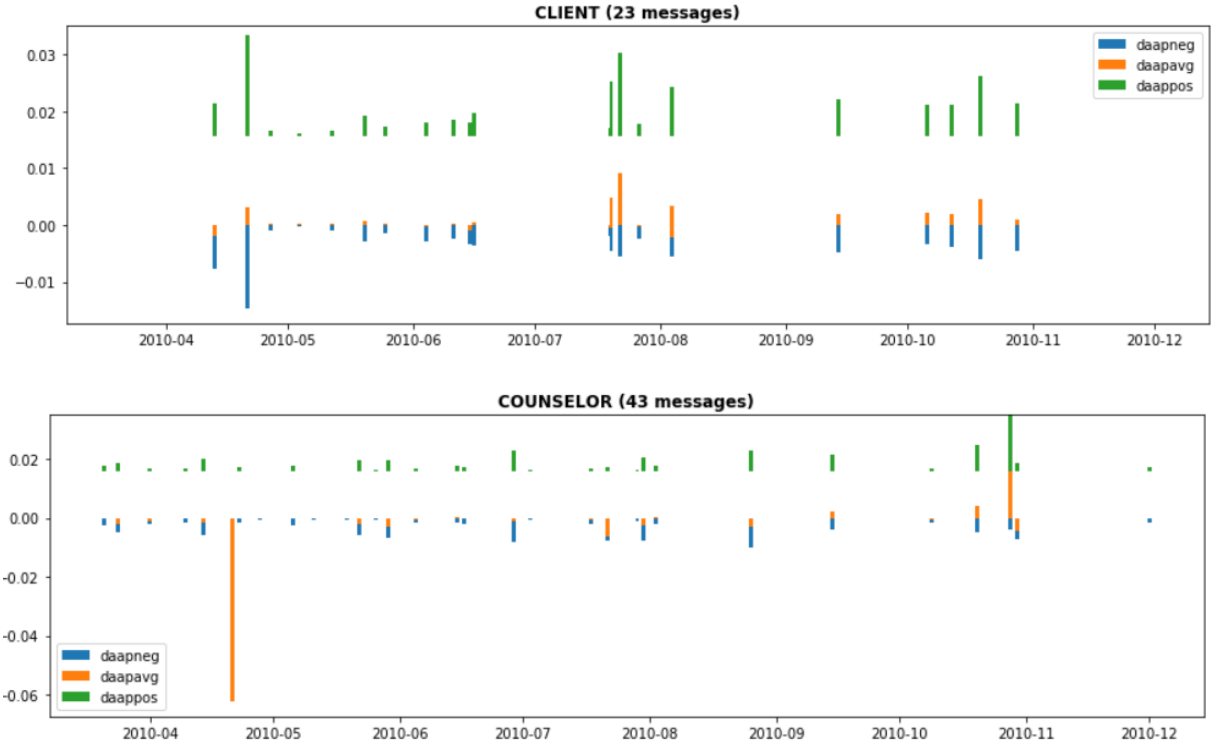
4


CLIENT (15 messages)
daapneg
daapavg
daappos
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
−0.0025
−0.0050
2010-03-12
2010-03-26
2010-04-09
2010-04-23
2010-05-07
2010-05-21
2010-06-04
2010-06-18
COUNSELOR (26 messages)
daapneg
daapavg
daappos
0.010
0.005
0.000
−0.005
2010-03-12
2010-03-26
2010-04-09
2010-04-23
2010-05-07
2010-05-21
2010-06-04
2010-06-18

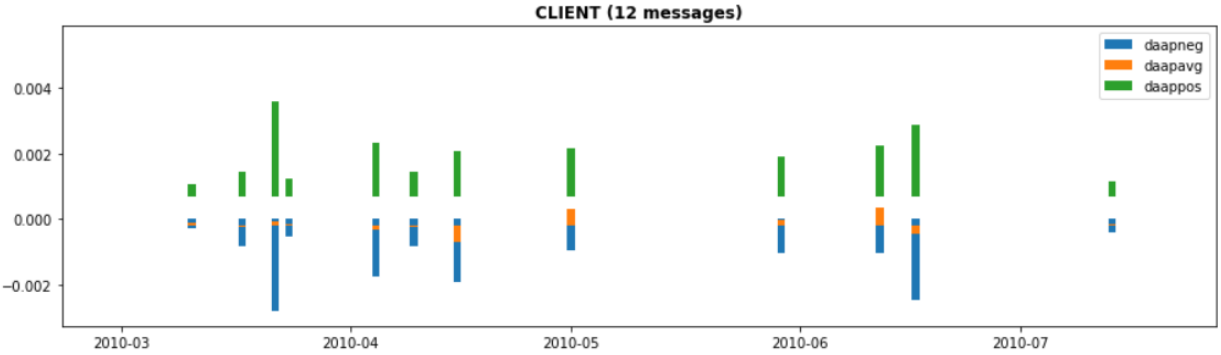
5

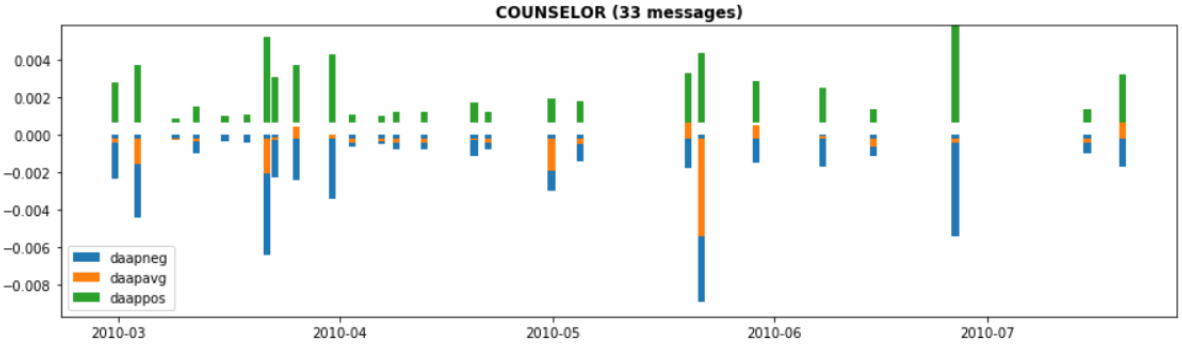

CLIENT (13 messages)
daapneg
daapavg
daappos
0.025
0.020
0.015
0.010
0.005
0.000
−0.005
−0.010
2010-03
2010-04
2010-05
2010-06
2010-07
2010-08
2010-09
COUNSELOR (35 messages)
daapneg
daapavg
daappos
0.025
0.020
0.015
0.010
0.005
0.000
−0.005
2010-03
2010-04
2010-05
2010-06
2010-07
2010-08
2010-09

6

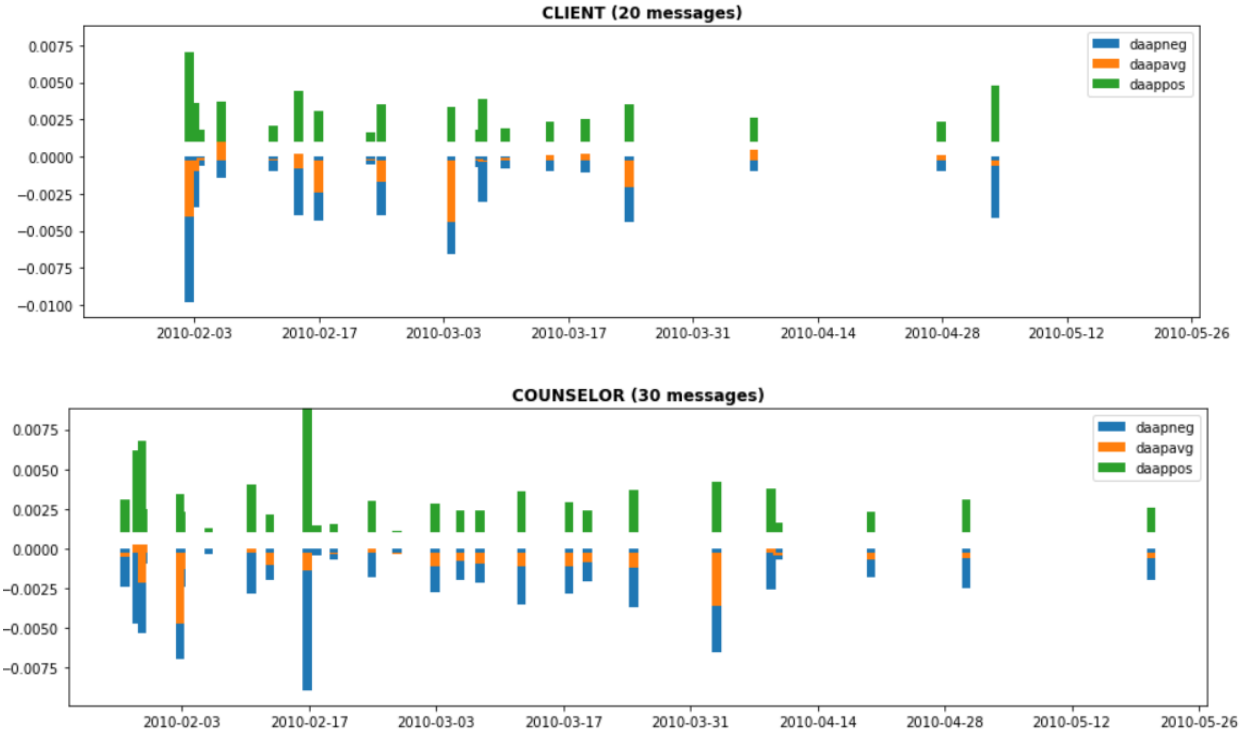

CLIENT (23 messages)
daapneg
daapavg
daappos
0.03
0.02
0.01
0.00
−0.01
2010-04
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10
2010-11
2010-12
COUNSELOR (43 messages)
0.02
0.00
-0.02
-0.04
-0.06
daapneg
daapavg
daappos
2010-04
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10
2010-11
2010-12

7

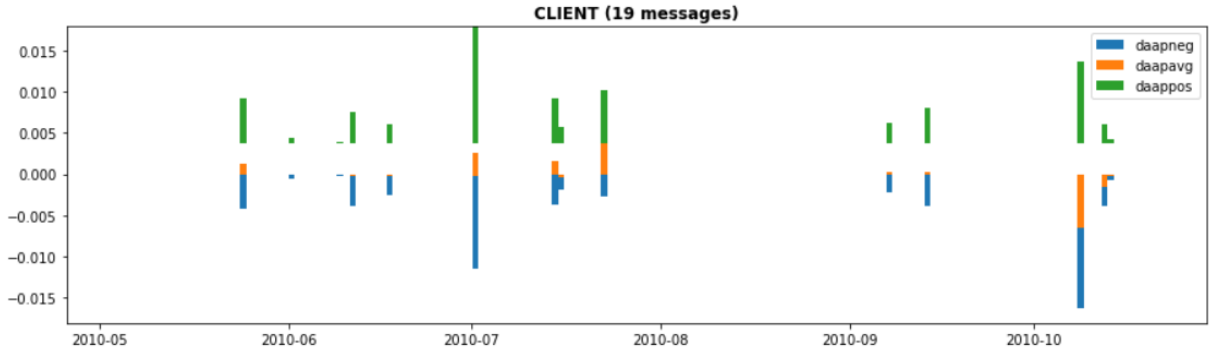

CLIENT (12 messages)
daapneg
daapavg
daappos
0.004
0.002
0.000
−0.002
2010-03
2010-04
2010-05
2010-06
2010-07

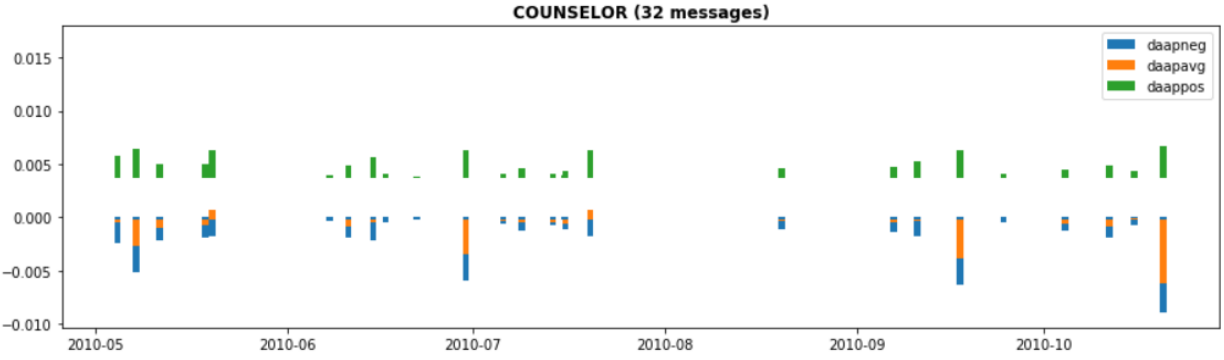

COUNSELOR (33 messages)
daapneg
daapavg
daappos
0.004
0.002
0.000
−0.002
−0.004
−0.006
−0.008
2010-03
2010-04
2010-05
2010-06
2010-07

8


CLIENT (20 messages)
daapneg
daapavg
daappos
0.0075
0.0050
0.0025
0.0000
−0.0025
−0.0050
−0.0075
−0.0100
2010-02-03
2010-02-17
2010-03-03
2010-03-17
2010-03-31
2010-04-14
2010-04-28
2010-05-12
2010-05-26
COUNSELOR (30 messages)
daapneg
daapavg
daappos
0.0075
0.0050
0.0025
0.0000
−0.0025
−0.0050
−0.0075
2010-02-03
2010-02-17
2010-03-03
2010-03-17
2010-03-31
2010-04-14
2010-04-28
2010-05-12
2010-05-26

9


CLIENT (19 messages)
daapneg
daapavg
daappos
0.015
0.010
0.005
0.000
−0.005
−0.010
−0.015
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10


COUNSELOR (32 messages)
daapneg
daapavg
daappos
0.015
0.010
0.005
0.000
−0.005
−0.010
2010-05
2010-06
2010-07
2010-08
2010-09
2010-10