Language Analysis of Motivational Interviewing Data

Master Thesis

Nicole Hollauf

ID: 1610310005

Klagenfurt, October 2018

Performed at the

Fachhochschule Kärnten/Carinthia University of Applied Sciences

Degree Program Health Care IT

in collaboration with

USC Institute for Creative Technologies

Supervised by

Dipl.-Ing. Dr. Michael Seger-Candussi | Dr. Stefan Scherer

Acknowledgements

I would like to express my deep gratitude to my supervisor Dr. Stefan Scherer who gave me the opportunity to write my Master Thesis in Los Angeles and I want to thank the Austrian Marshall Plan Foundation who made my stay abroad possible through their financial support. I also wish to express my sincere thanks to my supervisor Dipl.-Ing. Dr. Seger-Candussi for spending time providing useful suggestions about this thesis.

Finally, I wish to thank my parents for their support and encouragement throughout my study. Special thanks should also be given to all my new friends who inspired me and made my stay abroad unforgettable.

(Nicole Hollauf)

Abstract

In this work 187 motivational interviewing (MI) sessions from binge drinking patients have been examined for their patient and interventionist language. The patient language was categorized in MI typical categories follow neutral (FN), change talk (CT) and sustain talk (ST). The purpose of this work was to find the main topic of conversation and to investigate if there are significant differences between the sentiments and emotions from patients with post-session behavior change and patients without post-session behavior change. The topics of conversation were determined using topic modeling. The sentiments and emotions occurring in the patient and interventionist language were computed with sentiment analysis. Two methods for sentiment analysis were used: the NRC library and the LIWC dictionary. The Wilcoxon rank-sum test in the patient language category FN from NRC sentiments showed significant values in anger and negative emotions. In CT notable results were found in anticipation, positive emotions, joy, surprise and trust. LIWC scores were significant in pronouns, sadness and focus past from FN and in positive and negative emotions as well as focus past in CT.

Key words: motivational interviewing, sentiment analysis, topic modeling, patient language, binge drinking

Kurzfassung

In dieser Arbeit wurden 187 Interviews von Motivierender Gesprächsführung (MI) von Binge Drinker auf ihre Patienten- und Therapeutensprache untersucht. Die Patientensprache wurde in die MI-typischen Kategorien Follow Neutral (FN), Change Talk (CT) und Sustain Talk (ST) eingeteilt. Ziel dieser Arbeit war es, die Gesprächsthemen der Interviews zu finden und zu untersuchen, ob es signifikante Unterschiede zwischen den Gefühlen und Emotionen von Patienten mit einer Verhaltensänderung nach der Sitzung und Patienten ohne Verhaltensänderung nach der Sitzung gibt. Die Gesprächsthemen wurden mit Hilfe von Topic Model berechnet. Die in der Sprache auftretenden Gefühle und Emotionen wurden mittels Sentimentanalysen berechnet. Zwei Methoden zur Sentimentanalyse wurden verwendet: die NRC-Bibliothek und das LIWC-Wörterbuch. Der Wilcoxon-Rangsummentest in der Patientensprachkategorie FN mittels NRC zeigte signifikante Unterschiede zwischen Patienten mit Verhaltensänderung und Patienten ohne Verhaltensänderung in Wut und negativen Emotionen. In CT wurden nennenswerte Resultate in den Empfindungen Erwartung, positive Emotionen, Freude, Überraschung und Vertrauen gefunden. Die LIWC Ergebnisse waren signifikant bei Pronomen, Traurigkeit und Fokus Vergangenheit von FN und bei positiven sowie negativen Emotionen und Fokus Vergangenheit in CT.

Suchbegriffe: motivierende Gesprächsführung, Sentiment Analyse, Topic Modeling, Patientensprache, Rauschtrinken

Contents

Chapter 1

Introduction

According to a national survey from 2015, in the U.S. 15.1 million adults had Alcohol Use Disorder. Within these, 37.9 percent of college students participated in binge drinking [1]. Binge drinking among students is an omnipresent and leading problem on many U.S. college campuses and a risk factor for heavy drinking and alcohol dependence after college [2]. The National Institute on Alcohol Abuse and Alcoholism (NIAAA) defines binge drinking as a pattern of drinking that brings blood alcohol concentration (BAC) levels to 0.08 g/dl. Increase of blood alcohol concentration to a level of 0.08 g/dl or above typically occurs after four drinks for women and five drinks for men within 2 hours [3]. Beside the effects on health like fatalities, alcohol poisoning and hangover, binge drinking has also been linked to academic problems such as missed classes, reduction in class room performance, lower grades, dropping out and school failure [2]. An approach to combat alcohol problems such as binge drinking is called 'Motivational Interviewing'(MI).

1.1 Binge Drinking

In general a binge is described as a period of unrestrained, immoderate, excessive or uncontrolled self-indulgence [4].

Due to the wide range of research and changes in the intensity and frequency of drinking patterns several definitions of binge drinking have developed [5].

1.1.1 Definitions and Measures

1994 Wechsler and colleagues defined binge drinking as five or more drinks for men and four or more drinks for women on one or more occasions within two weeks prior their study [5]. This gender-specific measurements take the gender differences in the effects of alcohol consumption, including body mass and alcohol metabolism rates into account.

As already mentioned NIAAA defines binge drinking as a drinking pattern which increases the blood alcohol concentration levels to 0.08 g/dl or above. According to NIAAAm such an increase typically occurs after four drinks for women and after five drinks for men within 2 hours [3].

The definition of a drink is very vague. According to the World Health Organization (WHO), in the United States a standard drink is defined as a drink which contains up to 14 grams of pure alcohol [6]. In Austria a standard drink is a drink which contains 10 grams of alcohol [7]. However, the amount of liquid in the glass does not indicate how much alcohol is actually in the drink. Beverage serving sizes vary depending on the drink as well as the alcohol content within each type of beverage [6]. By definition, a standard drink in the U.S. corresponds to approximately 354 ml beer with 5 % alcohol content or 148 ml of red wine with 12 % alcohol.
Assuming a standard beer in Austria is 500 ml, makes this already two standard drinks and 125 ml red wine with 12 % 1.2 standard drinks based on WHOs definition [8].

Although the definition from NIAAA for the BAC cutoff of 0.08 g/dl is frequently used, this measurement also poses difficulties. People with a larger body mass or variation in the metabolism of alcohol could stay undetected even though they would be described as binge drinkers with other measurements [5].
The World Health Organization uses for the term binge drinking the term heavy episodic drinking and defines it as 60 grams or more of pure alcohol on at least one occasion in the past 30 days [9].
The Behavioral Risk Factor Surveillance System (BRFSS), which is a survey supported by Centers for Disease Control and Prevention (CDC), defines binge drinking as having 5 or more drinks for men and 4 or more drinks for women on one or more occasions in the past 30 days [10].
The Substance Abuse and Mental Health Services Administration (SAMHSA) defines binge drinking in their National Survey on Drug Use and Health (NSDUH) as five or more drinks on the same occasion on at least 1 day in the 30 days prior to the survey [11]. Table 1.1 shows different binge drinking definitions.

All in all, binge drinking definitions are similar and only vary in terms of the consumption time (e.g., on an occasion, within 2 hours) and period of binge drinking episodes (e.g., 2 weeks, 30 days) [5]. In general, binge drinking can be defined as X drinks for women and Y drinks for men in a time period of Z.

Questionnaires and screening tools are used to evaluate alcohol consumption. One tool to examine binge drinking behavior is Timeline Followback (TLFB). TLFB is a calender-based method where subjects estimate their daily drinking and mark days on which alcohol was consumed [5]. TLFB allows the estimation of the drinking behavior and captures drinking patterns. It incorporates daily versus weekend drinking, number of drinking days and drinking frequency [5].
The WHO developed the Alcohol Use Disorders Identification Test (AUDIT) to identify persons with harmful patterns of alcohol consumption [12]. The AUDIT is a screening method to capture excessive drinking behavior and to assist in brief assessment [12]. It consists of ten questions and a set of responses about recent alcohol use, alcohol dependence symptoms and alcohol-related problems [13]. Each response has a score ranging from 0 to 4, summing up to a maximum score of 40 [13]. Scores above 6 indicate at-risk drinking among young adults [5].

Table 1.1: Binge drinking definitions; BRFSS = Behavioral Risk Factor Surveillance System, CDC = Centers for Disease Control and Prevention, SAMHSA = Substance Abuse and Mental Health Services Administration [5]

NIAAA	Pattern of drinking alcohol that brings the BAC level to 0.08 g/dl or higher. This pattern corresponds to consuming 5 or more drinks for men,or 4 or more drinks for women in about 2 hours
WHO	Designated as Heavy episodic drinking - defined as 60 grams or more of pure alcohol on at least one occasion in the past 30 days
Wechsler et al.	Five or more drinks for men and four or more drinks for women on one or more occasions within two weeks
BRFSS	Consuming more than 5 drinks for men and more than 4 drinks for women on one or more occasion during the past 30 days
SAMHSA	Five or more drinks on the same occasion on at least 1 day in the past 30 days

The "binge score" is used to distinguish between binge drinkers and non-binge drinkers. The "binge score" bases on the Alcohol Use Questionnaire (AUQ) and is calculated from three specific AUQ questions (10,11,12) regarding drinking behavior [14]. A score under 16 categorizes non-binge drinkers while a score above 24 identifies binge drinkers [5].
The Composite Drinking Scale (CDS) is a score which combines four alcohol consumption measures [5]. The four consumption measures are the number of alcohol related occasions in the past 30 days, average number of drinks per week, number of drinks during partying and greatest number of drinks in one sitting during the last two weeks [5]. The CDS captures a range of relative risks and establishes a quadratic relationship between consumption and alcohol-related problems [5].
A modified version of the NIAAA six-item set of questions is the Alcohol Intake Questionnaire (AIQ) [5]. The AIQ provides qualifier for time and questions without time reference. Open-ended questions determine the frequency of binge drinking [5].

Another method to verify the classification of nondrinkers, moderate drinkers and binge drinkers are alcohol biomarkers. Regular excessive alcohol consumption alter alcohol biomarkers [5]. They can be classified in direct and indirect alcohol biomarkers [5].

Direct alcohol biomarkers are biomarkers which develop as a direct consequence of alcohol consumption. They are analytes of alcohol metabolism [15]. Analytes are created by nonoxidatively processes of the alcohol metabolism. Analytes can be measured for a longer period after alcohol consumption than alcohol itself [15]. Direct alcohol biomarkers are ethyl glucuronide (EtG), ethyl sulfate (EtS) and phosphatidylethanol (PEth).

Direct alcohol biomarkers can be found in blood, hair, nails and urine.

EtG and EtS exist in all body fluids and are usually measured in urine where it is detectable for 1 to 2 days [15]. However, EtG and EtS tests are very sensitive and the results can be interfered through daily use of alcohol containing products like mouth wash. More research regarding the effects of diseases, ethnicity, gender, genetic variation in enzyme systems on EtS and EtG is needed [15]. Furthermore, there are no reported relationships between different drinking patterns and the formation of EtG or EtS [5].

Phosphatidylethanol is a serum-based biomarker [15]. PEth is veritably 3 weeks after few days of moderate heavy drinking [15]. Piano et al. had shown, that there are significant correlations between levels of PEth and the total AUDIT score from young adult binge drinkers [5]. In addition, the sensitivity of PETh to alcohol was proven even with low doses (0.25 g/kg and 0.50 g/kg) of ethanol. However, more research is required to clarify the effects of binge drinking to biomarkers and health impairments [5].

Traditional alcohol biomarkers, also called indirect biomarkers, detect the toxic effects of alcohol to the body chemistry and organs. To the traditional biomarkers belong gamma glutamyl transferase (GGT), aspartate amino transferase (AST), alanine amino transferase (ALT), mean corpuscular volume (MCV) and carbohydratedeficient transferrin (CDT) [15]. GGT, AST and ALT are serum enzymes produced by the liver.

Alcohol consumption, liver damage and medication enhance the liver enzyme induction and cause increase in GGT [15]. GGT also increases with age and body mass index [5].
GGT biomarker show contradictory results regarding alcohol consumption. A study from Sillanaukee et al. showed a low to moderate but significant association of GGT levels and alcohol consumption. However, Meerkerk and colleagues, who investigated GGT levels dependency on different drinking patterns could not find predictive values to detect binge drinking [5]. A study from Pirro et al. found significantly higher GGT levels in heavy drinkers and distinguished heavy drinkers from social drinkers and nondrinkers. However, Piano et al. could not find differences in GGT levels among nondrinkers, moderate drinkers and binge drinkers [5].

Transferrin is a glycoprotein which is metabolized in the liver and relevant for iron transportation. Chronic and heavy alcohol consumption for about two weeks can cause transferrin to lose carbohydrate residues [5] [15]. CDT may also be affected by body mass index, female gender and smoking [5]. It is measured in the serum as the percentage of total transferrin that is carbohydrate deficient [15].
Piano et al. could detect substantial differences among heavy drinkerÂ´s, nondrinkers and social drinkers regarding CDT levels. Sillanaukee et al. also found a low to moderate but significant association of CDT levels with alcohol consumption [5]. According to Gonzalo et al. CDT shows high sensitivity and specificity in detecting chronic alcohol consumption greater than 60 g/day with no changes at levels lower than 30 g/day. However, Meerkerk et al. found low sensitivity and high specificity [5].

In detecting alcohol problems, CDT and GGT are interchangeable. The advantage of CDT is the insensibility of CDT regarding other factors than alcohol. However, CDT is also quite insensitive to episodic, heavy alcohol use which results in false negatives [5].

Elevations of the aminotransferases AST and ALT are indications of injury and liver cell death, which is often related to heavy drinking [15]. Pirro et al. compared AST levels from heavy drinkers, social drinkers and nondrinkers. They found a higher AST level in the group of heavy drinkers who consumed more than 60 g ethanol/day. ALT levels stayed the same. Nevertheless, there is a lack of studies investigating AST and ALT levels in dependency to binge drinking.

Mean corpuscular volume is the average volume of a red blood cell. Increased MCV values indicate drinking and macrocytosis. Further causes are older age, folate deficiency and gastrointestinal bleeding [5]. In consequences of all the possible sources of elevation, MCV is a very poor biomarker for alcohol consumption [15]. Studies from Piano et al. and Conigrave et al. do not show a significant increase of MCV among young adult binge drinkers [5].

1.1.2 Effects and Consequences of Binge Drinking

Since 1984 binge drinking has been a leading problem in U.S. universities. In 1984, more than 40 % of students were reported as participants in binge drinking in the United States. Due to the seriousness and the increased attention of the problem, the National Institute on Alcohol Abuse and Alcoholism founded a task force to develop a plan for binge drinking research at NIAAA [2].

The binge drinking pattern among students is regarded as temporary. Therefore, there are only a few studies that consider not only the short-term effects but also the long-term effects of heavy episodic drinking [2].

Binge drinking has health consequences but also affects people in the immediate environment [16]. According to NIAAA, annually estimated 88,000 people die from alcohol-related causes, which makes alcohol the third leading preventable cause of death in the United States after tobacco and poor diet [1]. In 2014 alcohol was the reason for 31 % of all driving fatalities.
Binge drinking is associated with an increased risk of short-term and long-term effects [17]. Short-term consequences are directly related to the intoxication. Hangovers, blackouts, memory loss, nausea and vomiting belong to the short-term effects of binge drinking. In extreme cases, binge drinking can lead to alcohol poisoning with occasional fatalities [17]. Amon students it is related to missed classes, lower grades and falling behind. In addition, binge drinking may be a factor in transmission of HIV and other sexually transmitted diseases. A meta-analysis revealed that the intention to engage unprotected intercourse increase by 5 % with a BAC rise from 0.1 g/ml [17].

In 1992 the Harvard School of Public Health College Alcohol Study (CAS) was initiated. The goal of the study was the national description of college student alcohol use and drinking behavior of this high-risk group [18]. CAS ran until 2006 and performed four surveys and more than 80 publications.

The main focus of CAS was on binge drinking, for the study defined as the consumption of five or more drinks for men and four or more drinks for women on one or more occasions during the two-week period prior the study. The study adapted the five-drink measure created by the University of Michigan and defined the gender-specific five/four measure as an indicator for heavy drinking [18].

The introduction of a gender-specific measurement took the gender differences in the effects of alcohol consumption, body mass and alcohol metabolism rates into account [18]. Furthermore, CAS additional measures were the consumption of alcohol within the last year, frequency of binge drinking (defined as drinking three or more times in the last 2 weeks), the number of drinking occasions in the past 30 days and the usual number of drinks at a drinking occasion [18].

The CAS study found a consistent national rate of binge drinking behavior among students of about 40 %. The polarization of drinking behavior between abstainers and frequent binge drinkers between 1993 and 2001 was protrude. 48 % of the drinkers stated, that the intention of their drinking was to get drunk [18].

According to CAS results, binge drinking at binge levels and beyond causes a series of problems in academic performance, social relationships, risk-taking behaviors and health related problems [18]. The academic performance problems include falling behind in schoolwork and lover grade point average. This problems are related to fewer hours which are spent on studying [18]. Regarding the social relationships, problems are reflected by antisocial behavior like vandalism or getting into trouble with the police when drinking [18]. The health risk problems include unplanned sexual activity and failure to use protection during sex and the resulting issues. According to Wechsler, half of the frequent binge drinkers, those who drunk at the five/four level or beyond three or more times in a 2-week period, experienced five or more different alcohol-related problems [18].

In context with drinking alcohol, one of the most dangerous side effects is intoxicated driving. According to NIAAA, 1,700 college students die per year from alcohol-related unintentional injuries, the majority in motor vehicle crashes [18]. 13 % of the students who drove regularly, reported to drive after consumption of five or more drinks. 23 % concedes that they rode with a driver who was intoxicated [18]. Binge drinking students are more likely to endanger themselves and others by operating or riding a motor vehicle after drinking [18].

Despite the experienced issues of frequent binge drinking students, less than 25 % contemplated themselves that they ever had an alcohol problem and only 13 % of this group thought they were heavy or problem drinkers [18].

The CAS study also incorporated secondhand effects of alcohol abuse. Secondhand effects included disruption of sleep or study, property damage and verbal, physical, or sexual violence. Three out of ten students reported experience in insult and humiliation by other drunk students and 19 % had serious arguments [18]. According to Hingson, 600,000 students per year experience violence or assault by intoxicated students. The sexual assault rate is higher at universities with higher binge drinking rates [18]. 5 % of female students were victims of sexual assault and 3 in 4 of these students were under alcohol influence. Furthermore, neighborhoods near universities with higher binge drinking rates than neighborhoods with a low rate were more likely to experience noise disruption, property damage and police visits [18].

1.1.3 Binge drinking influencing factors

In general, binge drinking behavior among students and the related research is a complex problem. Dowdall and Wechsler [19] extended their research with economic, political and ecological factors which are often neglected in other studies.

Figure 1.1 lists such factors. University campuses are often surrounded by bars and alcohol outlets with special promotions for students. Additional factors in Figure 1.1 are for example, campus environments such as fraternities or sororities where drinking plays a central role [19].

According to Dowdall, the investigated type of college or university is an important factor in college drinking behavior researches. Dowdall showed important differences in drinking patterns between women from women's colleges and women from co-educational colleges [19]. Furthermore, the location where the university or college is situated influences the drinking behavior. Alcohol availability, price as well as drinking tradition depends on the location [19].

As already mentioned, environmental factors play a significant role in study of binge drinking behavior.

CAS was designed to cover more than 100 colleges and thus the study includes environmental factors which may influence the drinking behavior [18].

Results of the study revealed a 1 % to 76 % variation of binge drinking behavior among the different colleges whereas the amount within the colleges itself remained the same. These findings indicate the importance of incorporation environmental factors. Binge drinking behavior also depends on the region of the country. In the northeastern and north-central states, the alcohol consumption rate is higher than in western states. Furthermore, the set of policies and laws governing alcohol sales and use influence the binge drinking behavior [18]. Wechsler clarifies that environmental factors such as residential setting, low price and high density of alcohol outlets as well as the prevailing drinking rates are related to the initiation [18]. Regarding the residential setting, binge drinking varies depending on the level of supervision. While students who live at their parents home show the lowest binge drinking rate, students who live off campus and in fraternity or sorority houses had the highest rates of binge drinking [18].

A factor which may reduce the binge drinking behavior of students is the demographic composition of a college [18]. A greater racial and ethnic diversity on a campus lowers the binge drinking rate among white majority students [18]. Furthermore, increased student

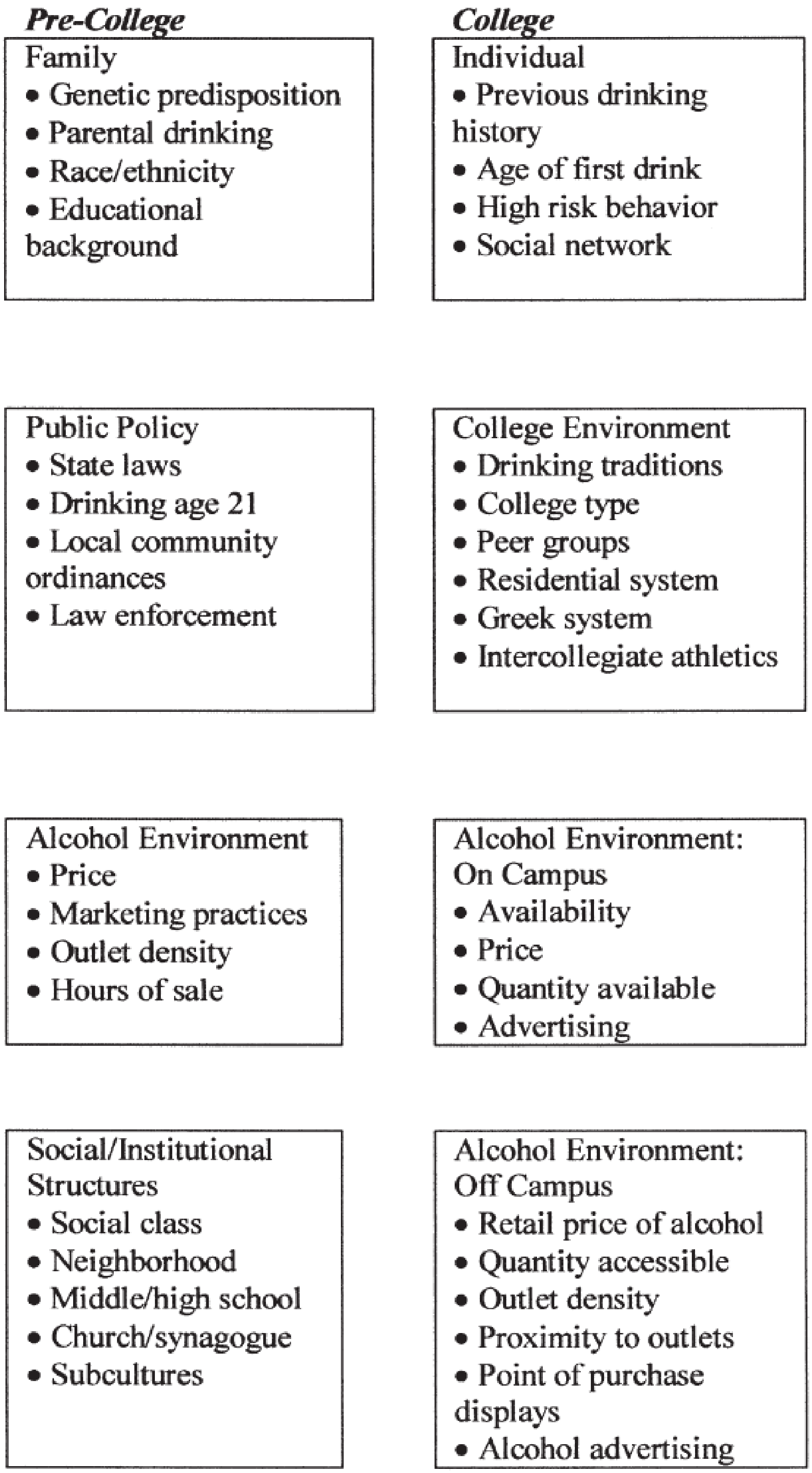

Figure 1.1: Factors affecting campus drinking [19]

participation in volunteer services limits the overall campus alcohol consumption and involved students were less likely to binge drink [18].

The alcohol price is also an meaningful factor for students drinking behavior. The cheaper and easier accessible the alcohol, the higher the binge drinking rate. Especially underage and female binge drinkers are sensitive to the price [18]. Another study discovered that the general binge drinking rate at colleges is higher in states which demonstrate a high binge drinking rate among adults [18].

1.2 Motivational Interviewing

Motivational interviewing (MI) is a counselling style that advises specific interpersonal and linguistic strategies to help clients identify and resolve ambivalences about behavioral changes [20]. Miller and Rollnick defined MI as a 'collaborative conversation style for strengthening a persons own motivation and commitment to change'. The goal of MI is to help people work through ambivalence and to commit behavior change [21]. It is often used to treat addictive behaviors like alcohol problems.

1.2.1 Counselling Style

MI is a counselling style whose method is composed of clinician style like warmth and empathy and technique like key questions and reflective listening [22].
Miller and Rollnick described five principles of MI [22] :

- Expressing empathy
- Developing discrepancy
- Avoiding argumentation
- Rolling with resistance
- Supporting self-efficacy

Classic approaches are more confrontational where therapists point out the need for change while the client denies it. Within MI, the therapist avoids confrontations and systematically attempts to elicit reasons for concern and change from the client itself. The therapist maintains a supportive atmosphere for exploration of ambivalent feelings. Resistance from the patient is not frontally confronted. It is redirected and focused back on open exploration. This process allows the patient to develop a motivational discrepancy between the present behavior and the desired goal behavior. [22]
A subliminal discomfiture arises which increases the probability to change the current behavior [23]. Classic addiction approaches are more aggressive and place the therapist in a powerful expert role. In MI the relationship between therapist and patient resembles more a partnership and acknowledges the clients personal responsibility and freedom of choice [22].

Most patients are conscious of their problem but they are ambivalent about the change. They want the change, but at the same time they do not want it [21]. They see reasons to change and reasons not to change. Ambivalence is willing both of two incompatible things. It is a normal human experience and ambivalence is an ordinary part of the change process [24]. People who do not want a change and are coerced by others, need to develop ambivalence about change as a first step into the direction of change [24]. As ambivalence is natural, ambivalent people have arguments for and against changing their current behavior. A patients argument to favor change is named *change talk* (CT)

[24]. Change talk is the clients desire, ability, reason and need to change [21]. A patients argument to remain the status quo and not changing the behavior is called *sustain talk* (ST). Sustain talk reflects the negative side of ambivalence and defends the current status [21]. Arguments which are neither positive or negative arguments about behavior change are called *follow neutral* (FN). A typical ambivalent sentence could be: "Yes I know that I need to stop drinking for my health (change talk) but I just need my evening beers to calm down (sustain talk). Arguments for and against change already reside within the ambivalent person [24]. Lecturing an ambivalent patient about change leads the person in the opposite direction and defending the opposite. This resistance and denial is also a normal nature of ambivalence [24].

As already mentioned, MI is a counselling style which uses key questions and reflective listening. MI requires a certain conversation dynamic. Experiments showed, that approaches containing convincing and persuading rise anger, defense, uncomfortableness and the feeling of powerless in the consulted person. However, when the consulted persons were asked person-centered key questions they felt engaged, empowered and understood. [24]

The theory of MI bases on interrelationships between clinician speech, client speech and client behavioral change [20], but the underlying mechanism is only partially understood [25].
It is assumed that behavior change is directly related to in-session CT and ST. Therapists using MI are able to influence client language and increase CT and decrease ST [26]. It is hypothesized that change talk is a mechanism of action in MI [25]. A meta-analysis found that MI-consistent (MICO; e.g. reflections, affirmations) and MI-inconsistent therapist behavior can elicit CT and reduce ST [26]. Therefore, CT is a mediator between MICO clinician behavior and improved client behavior change [25]. Since the solution to change is within the clients, they can talk themselves into change and only need redirection by their clinicians to trigger changes [25]. However, there are two therapeutic components which are proposed to increase the occurrence of CT. The technical component is the specific therapist behavior to elicit change talk. This component involves behaviors like reflections, open questions and affirmations. The second component is the relational component. The relational component focuses on global factors and the interrelationship between therapist and client such as therapist empathy and MI spirit [26]. In spite the reasons for CT as an indicator for post-session behavior change, there are studies which show contradictory results. Madson et al. showed, that CT itself did not indicate a behavior change, only in context to ST it was linked to outcomes [26]. Apodaca et al. extended this study and included the therapist measures MICO and MIIN. CT was not predictive of outcomes in this study nor were MICO or MIIN [26].

It might be possible, that CT is only a marker of other mechanisms of action in MI. Perhaps, highly motivated clients are more likely to offer change talk [25]. In that instance, CT would only be a signal for post-session behavior change rather than an actual cause of it [25].

1.2.2 MI Spirit

The underlying perspective with which one practices MI is called *the spirit of MI*. The spirit is the mindset and heartset of the practitioner [24]. Lack of spirit changes the MI session to a manipulation technique and its essence has been lost [27]. MI spirit is an essential component for the interpersonal relationship [27].
The MI spirit consists of four elements [24]:

Partnership The partnership aspects base on respect between clinician and patient - the counsellor creates a positive interpersonal atmosphere and skillfully guides the conversation

Acceptance The acceptance aspect is comprised of the four aspects worth, autonomy, empathy and affirmation

Compassion Compassion in MI is the active promotion of the others welfare and prioritizing the others needs

Evocation Within the evocation aspect the counsellor seeks to evoke and strengthen the patients motivation for change

All four elements consist further of an experiential and a behavioral component [24].
The key points of MI spirit are [27]:

The motivation for change is awakened in the client and not external imposed MI counsellors perform without coercion, persuasion, constructive confrontation or external contingencies

The clients have to resolve their ambivalence It is the client's task to articulate and resolve their ambivalence. The counsellor guides the client towards an acceptable resolution that triggers change

The counselling style is quiet and eliciting

Readiness to change is a fluctuating product of interpersonal interaction Resistance and denial are feedback for the therapist to modify his strategies

The relationship is like a partnership and not an expert/recipient role

Even though MI is represented more by the spirit rather than a technique, there are specific trainable therapist behaviors [27]:

- Understanding the person's frame of reference, especially through reflective listening
- Expressing acceptance and affirmation
- Eliciting and strengthening of the client's self-motivation
- Monitoring of the client's desire to change
- Affirming the client's freedom of choice and self-direction

1.2.3 Motivational Interviewing Coding Instruments

Motivational interviewing sessions are recorded, transcribed and assigned with specific codes. The basic coding units are utterances. Utterances are defined as a complete thought or a thought unit [28]. It is possible that more utterances appear one after the other without interruption. A sequence of utterances by on party is called volley [28]. Within the transcribing process the spoken words of the interview are written down and separated into utterances.

Motivational Interviewing Skill Code

The Motivational Interviewing Skill Code (MISC) was developed 1997. It is the original behavioral coding system and provides comprehensive information about the process and quality of MI [29]. Usage of MISC includes documentation of counsellor adherence, providing detailed session feedback, evaluation of the effectiveness of MI training, prediction treatment outcome and generation of new knowledge [30]. MISC is suited to investigate processes within MI sessions and provides detailed information about the behavior of the therapist and client during the MI session.

The MI session is coded in three separate passes. In the first pass, a global rating is assigned. In the second pass the counsellor behavior code is assigned and during the third pass the client behavior codes are added [30].

The global score is a ratio on a 7-point Likert scale about the global counsellor rating and the global client rating. The counsellor rating covers the three dimensions acceptance, empathy and spirit. It captures the rater's overall impression of the counsellors MI performance. The client rating reflects the client's highest level of self-exploration during the session [30].

In the second pass, the counsellor behavior codes are assigned. The behavior codes are determined based on categorization and decision rules. MISC 2.1 classifies the counsellor behavior in fifteen major categories [30] which are described in Table 1.2.

The third pass codes client utterances and assigns the client behavior codes. Client language is categorized in reason, taking steps, commitment and other. As already mentioned, client language which moves towards the desired target behavior change (TBC), is called change talk, language which moves away from the TBC is called sustain talk [30]. Utterances which do not indicate a movement towards or away from behavior change are called follow neutral. In order to identify ST and CT correctly, the clinician has to define the desired target behavior. The TBC is the desired target behavior e.g. 'Stop binge drinking'. Table 1.3 presents the client behavior codes.
Utterances which indicate change (CT) are coded with a positive (+) valence (e.g.: 'It is the right thing to do'(R+)). Utterances which reflect inclination away from the desired change (ST) are marked with a negative (-) valence (e.g.: 'It is the only way I can deal with the stress of my job' (R-))[30].

Table 1.2: MISC Counsellor Behavior Code

Code	Description
ADP/ADW	Advise with/without permission
AF	Affirm
CO	Confront
DI	Direct
EC	Emphasize Control
FA	Facilitate
FI	Filler
GI	Giving Information
QUC/QUO	Question closed/open
RCP/RCW	Raise Concern with/without permission
RES/REC	Reflect simple/complex
RF	Reframe
SU	Support
ST	Structure
WA	Warn

Sequential Code for Observing Process Exchanges

The Sequential Code for Observing Process Exchanges (SCOPE) was developed for the usage with recorded and transcribed motivational interviewing sessions [28]. SCOPE encodes interactions between therapist and client and focuses on sequential information between the parties. It investigates the relationships between theoretical MI constructs, general therapy process and client outcomes [31]. SCOPE is an adaption of MISC and the Commitment Language Coding System [31].

SCOPE is comprised of thirty therapists and sixteen client behaviors summing up in forty-six behaviors [32]. All utterances are coded with a behavioral code and in addition the therapeutic utterances are assigned with a quality code. The therapist behavior codes are presented in Table 1.4 [28].

Table 1.3: MISC Client Behavior Code

Code	Description	Subcategory
R	Reason	subcategories : desire (d), ability (a), need (n)
O	Other	
TS	Taking Steps	
C	Commitment	
FN	Follow Neutral	

Each of the therapist utterances is additionally denoted with the quality code M+, M- or M0 [28]. The quality code represents the degree of ideal MI practice [32]. For example, M+ is assigned if the statement expresses empathy, develop discrepancy, support self-efficacy or minimize resistance [28].

Within SCOPE, the client utterances are categorized in one of the three categories ask, follow/neutral (FN) or commitment language (change talk, sustain talk). Commitment language is further subclassified in one of the following classes described in Table 1.5 [28]. Each of this subclasses is denoted with + or -, depending on the movement towards or away from the target behavior change [28].

Motivational Interviewing Treatment Integrity

The Motivational Interviewing Treatment Integrity (MITI) is derived from the MISC. It was developed to create a less complex rating system which focuses on therapist functioning [33]. With the assistance of MITI the clinician can be evaluated [34]. It is also used for clinician training and quality check in clinical trials [35]. MITI evaluates MI components inclusive engaging, focusing, evoking and planning [34].

MITI is comprised of two components, the global score and the behavior count [34]. The global score is assigned from the rater depending on their overall impression of the clinicians MI skills. It reflects a holistic evaluation of the interviewer. The global score is a single number on a five-point Likert scale between 1 (minimum) and 5 (maximum) assigned to each of the four global dimensions. The four global dimensions are cultivating change talk, softening sustain talk, partnership and empathy [34].

Cultivating change talk Measurement to which extent the clinician actively encourages the patients change talk - low scores are a result of clinicians inattentiveness about change talk

Softening sustain talk Measurement to which extent the clinician avoids focusing at sustain talk - absence of sustain talks attains a high score at this scale

Table 1.4: SCOPE Counsellor Behavior Code

Code	Description	Subcategory
Adv	Advise	coded as inform,direct or question
Aff	Affirm	
Con	Confront	
Dir	Direct	
Econ	Emphasize Control	
FB	Feedback	
Fill	Filler	
Sdis	Self-Disclose	
GI	General Information	
Perm	Permission seeking	
CQ/OQ	Closed/open Question	
RC	Raise Concern	
SR/CR	Simple/complex Reflection	
Sup	Support	coded as support or confront
WA	Warn	

Table 1.5: SCOPE Client Behavior Code

Code	Description
C	Commitment
D	Desire
A	Ability
N	Need
R	Reason
TS	Taking Steps
O	Other

Partnership This scale measures the clinicians understanding that changes resides within the patient - clinicians who create an interview with a two equal partner atmosphere are high on this scale

Empathy Measures the clinicians effort to understand the patients perspective

The behavior count is a counter of particular clinician behaviors. The behavior count captures clinician behavior based on categorization and decision rules, rather than the overall impression [34]. The clinicians volleys are coded with the following codes:

- Giving information (G)
- Persuade (Persuade)
- Question (Q)
- Reflection simple/complex (SR/CR)
- Affirm (AF)
- Seeking collaboration (Seek)
- Emphasizing autonomy (Emphasize)
- Confront (Confront)

Each volley can be assigned with a maximum of eight codes. Only one of the eight codes may be assigned [34].
MITI reviews a random 20-minute segment of the entire MI session. Shorter or longer segments are not beneficial. The segment randomness should be given [34].

Other coding systems

In research settings MI coding is used for monitoring internal validity of clinical trials, evaluation of the effectiveness of MI training, understanding the relationship between clinician and client and to model the specificity of post-session behavioral change in dependency of in-session behavior prediction [20]. Therefore behavioral coding is an essential tool in order to provide feedback and maintain high-quality in MI practice.
Therapy sessions are often coded manually with human raters. Manual methods require training and are often time- and cost-intensive [20]. For example, a single 50-minute session can have 12,000 to 15,000 words within several hundred utterances and coding can take between several hours to code in full, depending on the complexity of the coding system [20]. In research questions, a study may consists of hundreds of MI-sessions resulting in an substantial expense of time, money and personnel resources.
Using natural language processing and machine learning techniques, rather than human coders, to code MI sessions delivers promising results [20].

Tanana et al. developed two natural language processing models to assign utterances with the MISC code. As an input variable, both models used dependency trees. Dependency trees are grammatical structures which link words into a hierarchical structure based on their order and relationship between words [20].

The first model is called *Discrete Sentence Feature (DSF) Model*. It bases on N-grams and dependency trees as an input to the prediction model. The relative likelihood of an utterance falling into a specific MISC category was predicted by presence or absence of sentence features within an utterance [20].

The second natural language processing model is called *Recursive Neural Network (RNN) Model*. This model is based on recursive neural networks. Within the RNN model, each word in an utterance is assigned a vector of 50 numeric values which represent the semantics of each word within a 50-dimensional latent space [20]. The numeric vector represents semantic similarities between words and semantic differences respectively. All vectors from the words of one utterance are combined to a single output with values between zero and one [20]. These vectors are used as an input to a multinomial regression model to predict the MISC codes.
Both models created by Tanana et al. predicted most of the MISC codes comparable to manually coding. However, not every MISC code was predicted correctly.

Another method which codes client change talk was developed by Huang et al. Huang et al. created a neural network to automatically code client change language named *Contextual Hierarchical Attention-based Recurrent Model (CHARM)* [26]. In addition to the automatic coding of utterances, CHARM captures patterns which are predictive for patient outcomes.
In contrast to the methods mentioned before, CHARM includes the verbal behavior of the patient and interventionist as well as dyadic codes to predict patient language codings (CT, ST, FN) [26]. CHARM utilizes three kinds of information to assess utterances: preceding interventionist verbal behavior (context), prior MISC annotations of utterances (codes) and the current patient utterance (content) [26]. Each of these information is composed of a bidirectional recurrent layer, an attention layer and a composition layer. The bidirectional recurrent layer encodes words in each utterance. The attention layer is used to estimate the importance of the words for the utterance. The composition layer is a vector representing current patient utterance, the proceeding interventionist context and the prior assigned MISC codes [26].
The results showed that CHARM resulted in a robust method which has improvements compared to other methods which ignore the therapists language behavior.

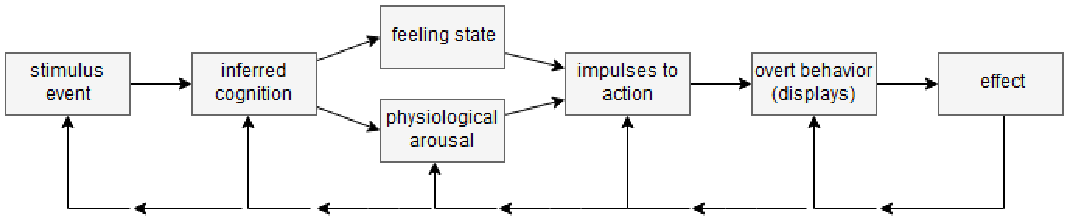

Figure 1.2: Components of the emotional feedback loops [36]

1.3 Emotions and Sentiments

Due to the focus of the work on sentiment and emotion analysis, the nature of emotions is described below. Because of the complexity of this psychological topic, only an overview is given.

The nature of emotions is a complex psychological topic. According to Plutchik, the function of emotions is to restore the individual state of equilibrium when unexpected or unusual events create disequilibrium [36]. He compares emotions to a chain of events made up of feedback loops. External and internal stimulus events are primary triggers which start the emotion process. In general, cognitions mark the start of the chain. However, through a feedback process they can be influenced by events which appear later in the chain. Such events may be arousal or ego defense [36]. Figure 1.2 presents the components of the feedback loop. The process starts with a stimulus event followed by the inferred cognition. The sensory information is evaluated and results in an action which normalizes the relationship between the individual an the triggering event [36]. A more detailed overview is given in Figure 1.3 which shows an example of a feedback loop for the emotion of fear.
A threat is recognized as danger resulting in fear and increased autonomic activity which triggers the impulse to flee. Fleeing may resulting in the reduction of threat and therefore to the normalization of the relationship between the individual an the triggering event [36].
Plutchik defines emotions as a behavioral homeostatic, negative-feedback system. Emotions are a homeostatic process in which behavior mediates progress toward equilibrium [36]. The baseline of this description is formed by the evolutionary perspective, that emotions have a function in the lives of individuals.

Due to hundreds of emotion words which occur in similar families, Plutchik conceptualized the primary emotions analogous to a color wheel. His concept is known as Plutchik's Wheel of Emotions which is represented in Figure 1.4. Plutchik's Wheel of Emotions places similar emotions close together and their opposites 180 degree apart, comparable to primary colors and complementary colors [36]. Mixture of emotions create other emotions (cf. mixing of colors). Plutchik extended his concept with a third dimension representing the intensity of emotions, resulting in a cone shaped structural model [36]. The circle of the wheel represents the degree of similarity among emotions. The vertical dimension of the cone represents the intensity of the emotion.

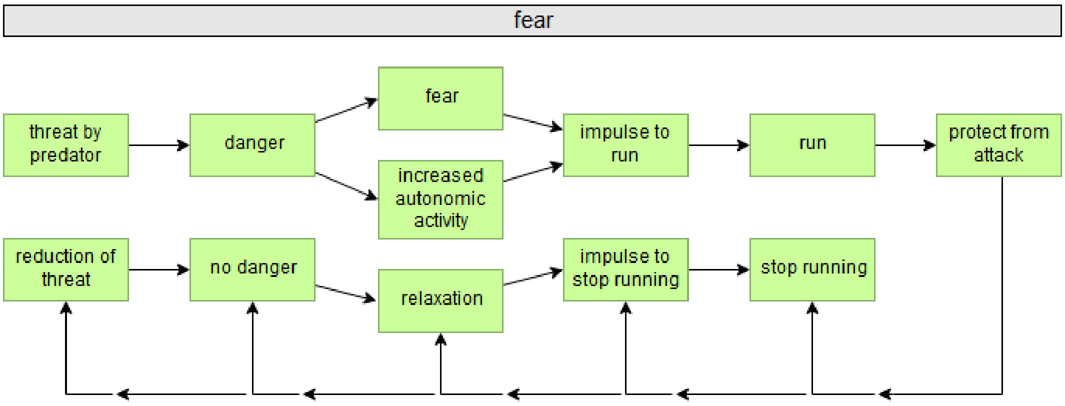

Figure 1.3: Example feedback loop for the emotion of fear [36]

The psychoevolutionary theory assumes eight basic emotion dimensions which are represented by the eight sectors of Plutchik's Wheel of Emotions. The eight basic emotions and therefore the primary emotions of the wheel are joy, trust, fear, surprise, sadness, disgust, anger and anticipation [36]. The emotions in the blank spaces are mixtures of two primary emotions also called primary dyads. Such dyads are for example the mixture of joy and acceptance which produces the emotion of love. The mixture of disgust and anger results in hatred or hostility.

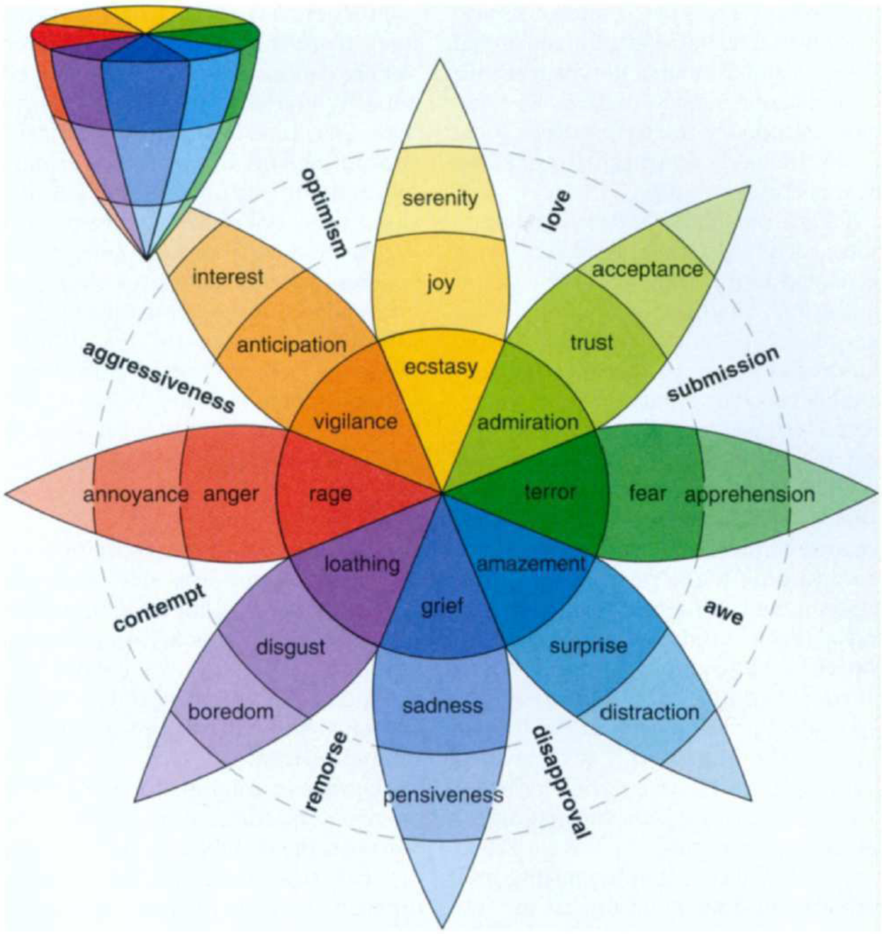

Figure 1.4: Plutchik's Wheel of Emotions [36]

According to Ben-Ze'ev emotions consists of the two dimensions feeling and intentionality and can be described at three levels: neurophysiological-biochemical level, motor or behavioral level and the mental level [37]. The first level is related to neurophysiological-biochemical body functions like neurotransmitters as well as the autonomic and somatic nervous system. The second level describes body reactions like facial coloring as well as various types of behavior. The mental level refers to the basic dimensions feeling and intentionality. Intentionality is comprised of the components evaluation, cognition and motivation [37]. The mental component is described below.

Emotions are evaluative attitudes. The emotional attitude presupposes a certain norm, the fulfilment or deviation of which is the basis of the emotional evaluation [37]. For example, hating someone implies a negative evaluation of the person. Emotional evaluation can be distinguished between deliberate and undeliberate evaluation. Deliberate evaluation is a conscious thought process, undeliberate evaluation is based on the primitive and basic evaluation called instinct [37].

Having an emotional attitude to something requires information (truthful or distorted) about it [37]. This fact is represented by the cognition component. For example, the fear of riding a motorcycle presupposes information about motorcycles and the evaluation of them as dangerous [37].

The desire to maintain or change a current situation is profound in the motivational component. Emotions are close connected to the motivational realm and the activity of the organism [37]. Specific actions for example, are often justified and explained by referring to emotions. In an emotion such as anger, this motivation is typically expressed in overt behavior [37].

The feeling dimension is a mode of consciousness connected with the persons own state. The feeling dimension has the lowest level of awareness and does not have a meaningful, cognitive content. It is a mode of awareness. The feeling dimension includes feelings like thirst, hunger, pain, thrill or sleepiness [37]. Feelings are a central feature of emotions and can be defined as an 'emotional attitude'. Emotions then again, can be defined by referring to feelings (e.g. feeling emotions, 'I feel ashamed').

Chapter 2

Methods

2.1 Knowledge Discovery and Data Mining

The amount of generated text increased dramatically in the recent years. In order to deal with all the knowledge and information of these texts it is necessary to find a way to process this data [38]. Knowledge discovery is the overall process of discovering useful knowledge from data.

Data mining is the application of particular algorithms to extract patterns and therefore a specific step in the knowledge discovery process [38].

According to the Cross Industry Standard Process for Data Mining (CRISP-DM) Methodology there are six main stages in the data mining process [39] [38] :

1. **Determine objectives** Determination of the application and required data as well as identification of the goal of the overlying knowledge discovery process
2. **Data understanding** Data acquisition and description of the involved data
3. **Data preparation** Data cleaning and preprocessing
4. **Modeling** Selecting modelling technique, model building and model assessment - modeling is an iterative process where parameters are changed until the optimal model is found
5. **Evaluation** Assessment of the data mining results and approving the models
6. **Deployment** Determine a strategy for result deployment

2.1.1 Text Mining

Text Mining is an analytical method to extract semantic structures and meaningful information out of unstructured data like text data. It is part of knowledge discovery and data mining [38]. Text mining covers a wide range of topics and algorithms such as [38] :

- Information Retrieval (IR)
- Natural Language Processing (NLP)
- Information Extraction from Text (IE)
- Text Summarization
- Unsupervised Learning Methods
- Supervised Learning Methods
- Text Streams and Social Media Mining
- Opinion Mining and Sentiment Analysis
- Biomedical Text Mining

Text mining presupposes a specific data structure to facility the document analysis [38]. A common approach is a vector representation which contains the number of occurrences of each term without taking the word order into account. This approach is also known as bag-of-words (BOW) [38]. The resulting vector can be analyzed with dimension reduction techniques. The main dimension reduction techniques used in text mining are Latent Semantic Indexing (LSI), Probabilistic Latent Semantic Indexing (PLSI) and topic models [38].

2.1.2 Topic Modeling

Topic Modeling is a text mining method and a convenient approach to analyze large amounts of unclassified text and helps to identify patterns of words in documents [40]. Topic models are based on hierarchical probabilistic models and can be generalized to any other kind of data such as images, biological data or survey information [40]. Topic models find patterns of word occurrence and connect different documents which share the same pattern.

Topic models assume that each document is a mixture of specific topics [41]. These models assume a set of latent topics. Latent topics are multinomial distributions over words and topic models assume that each document can be described as a mixture of these topics [41]. The topics probabilities represent the documents [42]. In text analysis, topic models base on the bag-of-words assumption which means that the order of words in a document can be neglected [42]. The bag-of-words assumption is, in the probability theory, also known as "exchangeability".
Each document can be represented by histograms which contain the occurrence of words [43]. The histograms are a distribution over a certain number of topics and each topic is a distribution over the words in the vocabulary [43].

There are several kinds of topic modeling approaches but all of these topic models build the same type of latent space. They build a topic collection for the corpus and a collection for topic proportions for each of its documents [41].

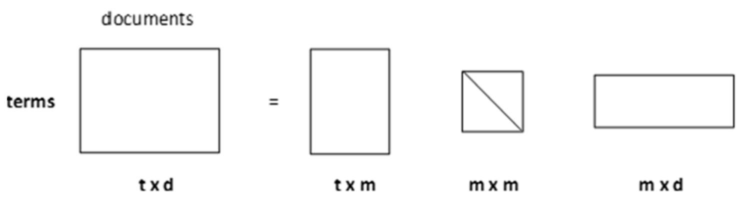

Figure 2.1: Schematic representation of SVD where t denotes the terms, d the documents and m the user defined matrix [47]

Latent Semantic Analysis

Latent Semantic Analysis (LSA) former, Latent Semantic Indexing, was originally developed for Information Retrieval. Latent Semantic Analysis is a vector based method to map the high-dimensional vectors to a lower dimensional representation, the so called *latent semantic space* [44]. The goal is to find a data mapping which represents the semantic relation between words and documents in relation to their proximity in semantic space [44].

In regards of LSA for text analysis, text is represented as a matrix in which each row stands for a unique word and each column stands for a text passage. Each cell contains the number of the words in the specific text passage [45]. Mathematically, a document collection $D = \{d_1, ..., d_n\}$ with words $W = \{w_1, ..., w_m\}$ can be represented with the co-occurrence matrix $N \times M$ with $N = (n(d_i, wj))_i j$, where $n(di, wj)$ is the frequency of the word w_j in the document d_i [44]. N is called term-document matrix with the row/column as the document/term vectors [44]. Like other topic models assumptions, LSA uses the bag-of-words approach. One problem of LSA is data sparseness which means, that the likelihood to find common terms in related documents may be small when the documents do not contain the exact same words [44].

As already mentioned, the goal of LSA is the mapping of documents to the latent semantic space. The latent semantic space has typically the order of $\approx 100 - 300$ dimensions. The mapping of the co-occurrence matrix is realised with Singular-Value Decomposition (SVD). It is a matrix algebra technique to re-orientate and rank dimensions in a vector space [46], see Figure 2.1. SVD decomposes rectangular matrices into the product of three other matrices [45]. SVD allows meaningful association values between document pairs, even if the documents do not have common terms contrary to LSA without SVD [44].

Probabilistic Latent Semantic Analysis

Probabilistic Latent Semantic Analysis (PLSA) is a statistical latent class model which is more suitable for term matching than LSA [48]. PLSA uses a latent variable z which represents a class or topic. The conditional probability between documents d and words w is modeled through the variable z [48]. The PLSA model is denoted by $P(w|z)$ and $P(z|d)$. $P(w|z)$ is the probability of words in a given class and it is conditionally independent of the document. Words w can belong to more than one class and a document d can describe more than one topic [48]. The joint probability of a word d and a document d can be described with Equation 2.1 :

$$P(w, d) = P(d) \sum_{z} P(w|z)P(z|d) \tag{2.1}$$

The parameters $P(z|d)$ and $P(w|z)$ are estimated using the iterative Expectation-Maximization (EM) algorithm [48]. EM bases on discovering of the maximum likelihood estimation of parameters when the data model depends on certain latent variables [49]. A drawback of PSLA is, that PSLA does not provide a probabilistic model at the document level [42]. The documents are represented as a list of numbers which leads to two problems [42] :

1. The number of parameters grows linearly with the size of the collection
2. It is uncertain how to assign a document a probability outside of the training set

Latent Dirichlet Allocation

Latent Dirichlet Allocation (LDA) is a generative probabilistic model for collections of discrete data such as a collection of documents also known as text corpora [42]. LDA is one of the most popular topic modeling methods to analyze a large amount of data [47]. LDA is a three-level hierarchical Bayesian model and presupposes the bag-of-words approach. Each item in a collection is modeled as a finite mixture over an underlying set of topics and each topic is modeled as an infinite mixture over an underlying set of topic probabilities [42]. According to de Finetti, any collection of exchangeable random variables can be represented as a mixture distribution - in general an infinite mixture [42].

The idea of LDA is the representation of documents as random mixtures over latent topics and each topic is described by a distribution over words [42]. Each document consists of multiple numbers of topics with a different percentage [47]. For example, a document might be comprised mainly of words from Topic A and only to a small percentage of words belonging to Topic B [47].

As already mentioned, LDA is a generative probabilistic model. The process is described as followed [47] :

1. For $k = 1$ to K :
   (a) $\Phi^{(k)} \sim Dirichlet(\beta)$
2. For each of the M documents in dataset$D$:
   (a) $\theta_{(d)} \sim Dirichlet(\alpha)$
   (b) For each word in document d :
      i. $z_i \sim Discrete(\theta_{(d)})$

ii. $w_i \sim Discrete(\Phi^{(z_i)})$

The number of topics is a free parameter K in the data D where k is an instance of a topic, d an instance of a document and w an instance of a word [47].
$\Phi^{(k)}$ is the probability distribution form topic k and $\theta_{(d)}$ is the probability distribution of document d over all topics K. A topic z_i is assigned to each word w_i. α and β are parameters of the Dirichlet distribution. $\Phi^{(w_i)}$ is the probability distribution of topic assignment for word w_i [47]. These parameters are approximations. *Dirichlet* are draws from a uniform Dirichlet distribution with scaling and *Disecrete* are draws from these parameters. After the selection of K topics k, LDA assigns each word of the documents to one topic based on a Dirichlet distribution.

The generative process is described with Equation 2.2 [47] :

$$p(w_i, z_i, \theta_d, \Phi^{(k)}|\alpha, \beta) = p(\Phi|\beta)p(\theta|\alpha)p(z|\theta)p(w|\Phi_z) \tag{2.2}$$

z_i,θ_d and $\Phi^{(k)}$ are latent variables. LDA uses this variables to classify the words in the data to a specific topic. The probability of the words in a document is calculated with Equation 2.3 [47] :

$$p(w_i|\alpha, \beta) = \int p(\theta_d|\alpha) \prod_{i=1}^{I} \sum_{z_i} p(z_i|\theta_d)p(w_i|z_i, \beta)d\theta_d \tag{2.3}$$

The probability of the documents in the data D is shown in Equation 2.4 [47] [42] :

$$p(D|\alpha, \beta) = prod_{d=1}^{M} \int p(\theta_d|\alpha)\Big(\prod_{i=1}^{I_d} \sum_{z_{di}} p(z_{di}|\theta_d))p(w_{di}|z_{di}, \beta)\Big)d\theta_d \tag{2.4}$$

A schematic description of the Latent Dirichlet Allocation is shown in Figure 2.2 [42].

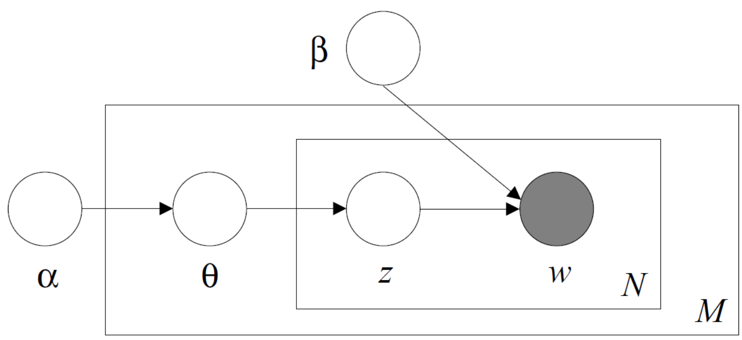

Figure 2.2: Schematic representation of LDA where the outer box M represents the documents and the inner box N denotes the selected topics and words of the document constrained by the parameters [47] [42]

2.2 Sentiment Analysis

Sentiment analysis, also known as opinion mining is the analysis of people's opinions, sentiments, attitudes and emotions from written language. Due to the growth of social media and the associated interest in opinions, sentiment analysis is an active research area in natural language processing [50]. Opinions are central to all human activities and have a significant influence on our behavior, therefore sentiment analysis systems are widely used in corporate and social areas [50]. Sentiment analysis can be classified in three main levels:

Document level At the document level, sentiment analysis inspects the whole document if its overall sentiment is whether positive or negative (e.g., for a product review). The document level assumes that each document expresses positive and negative opinions towards a single entity (e.g., a single product) [50]

Sentence level The sentence level analyses whether a sentence expresses a positive, negative or neutral opinion, where neutral means no opinion [50]

Entity and aspect level The entity and aspect level is a finer-grated analysis than the analysis on document or sentence level. It determines what people like and what people do not like. Aspect level is based on the idea that an opinion consists of a sentiment (positive or negative) and a target of the opinion. In many applications opinion goals are described by entities and their various aspects. The aim of this analysis is to discover sentiments about entities and their aspects [50]

Indicators of sentiments are called sentiment words. Sentiment words are words which are used to express a positive or negative emotion. Automatically analyzing text and detecting emotions such as joy, sadness, fear, anger and surprise is applied in many areas like identifying blogs which express specific emotions towards a topic, identifying newspaper headlines or for developing automatic dialog systems [51]. Different emotions are expressed through different words. For example, *good* and *amazing* are positive sentiment words, *delightful* and *yummy* emphasize joy and *cry* or *gloomy* express sadness. A collection of such words and emotions is called sentiment lexicon or emotion lexicon [51]. Such lexicons are useful for automatic methods to identify emotions evoked by a word, even when words may evoke different emotions in different contexts [51].

2.2.1 Sentiment Analysis with RStudio

As mentioned above, lists which contain sentiment words are called sentiment lexicon. The RStudio package *Tidy Data* from [52] contains three general purpose lexicons:

- NRC from [53]
- bing from [54]
- AFINN from [55]

These lexicons are based on single words, also called unigrams. In natural language processing n-grams are a n-long sequence of text. A unigram for example is a single word like *happy*, a bigram would be *not happy* and a trigram *not happy because*. The total sentiment of a document is calculated by summing up the number of individual sentiments for each word in the document [56]. The three lexicons are described as followed [56] :

NRC The NRC lexicon contains the eight basic emotions joy, anger, anticipation, disgust, fear, sadness, surprise and trust and the two sentiments negative and positive

bing The bing lexicon categorizes words into positive and negative sentiments

AFFIN The AFINN lexicon is score based and assigns words a score between -5 to 5, referring negative sentiment words to a negative score and positive sentiment words to a positive score

In this work the NRC (NRC stands for National Research Council Canada) lexicon was used. The NRC lexicon was created by Saif M. Mohammad and Peter D. Turney. The lexicon contains 14,182 unigrams (words) which results in approximately 25,000 word senses and is available in over hundred languages.

The workflow of the executed sentiment analysis is based on [56]. The first step was to bring the data into a tidy format, which means that tokens had to be created. The tokens represent the text to be analyzed as one-word-per-row format. When the data is in the tidy format, the sentiment analysis is done as an *inner_join()* operation from the R package *dyplr* [57] [56]. The *inner_join()* operation returns all rows from x where there are matching values in y, and all columns from x and y. If there are multiple matches between x and y, all combinations of the matches are returned [57]. In this sentiment analysis x is the generated tidy data, also denoted as tokens, from the utterances from Changer and Non-Changer. After tidying the data, stopwords were removed with an *anti_join()* operation. The *anti_join()* operation returns all rows from x (tokens) where there are no matching values in y (stopword list), keeping only the columns from x [57].

2.2.2 Linguistic Inquiry and Word Count

Linguistic Inquiry and Word Count (LIWC) is a dictionary and text analysis application. The current version is called LIWC2015. LIWC2015 is used as a method for studying various emotional, cognitive and structural components occurring in written speech [58]. The LIWC2015 dictionary is based on unigrams.

LIWC2015 accesses single text files, file groups or texts within a spreadsheet. Each file is analyzed sequentially. LIWC2015 analyzes each word and searches for a match in its dictionary file. If a match is found, the counter (score) for the appropriate word category is incremented [58].

The LIWC2015 dictionary contains 6,400 words, word stems as well as emotions and each entry additionally defines one or more word categories or subdictionaries [58]. As an example, the word "cried" is part of five categories. The five categories are sadness, negative emotion, overall effect, verbs and focus past. If the analyzed text contains the word cried, the score for each of this five categories is incremented [58]. LIWC2015 is also capable to recognize word stems. For example it contains the word *hungr**. Every word in the text starting with *hungr* is counted as an ingestion word. *Hungr* may stand for hungry, hungrier or hungriest. The asterisk in the dictionary indicates the acceptance of all following letters, hyphens or numbers in this word [58]. In this work, twenty categories of the LIWC2015 dictionary were used. The twenty categories with its abbreviations and example words are described in Table 2.1. A complete list of the LIWC2015 word categories can be found in [58].

Table 2.1: Overview of the used LIWC2015 categories

Category	Abbreviation	Example	Amount
1st pers singular	i	I, me, mine	24
1st pers plural	we	we, us, our	12
3rd pers singular	shehe	she, her, him	17
3rd pers plural	they	they, their, theyâ€™d	11
Total pronouns	pronoun	I, them, itself	153
Positive emotion	posemo	love, nice, sweet	620
Negative emotion	negemo	hurt, ugly, nasty	744
Anxiety	anx	worried, fearful	116
Anger	anger	hate, kill, annoyed	230
Sadness	sad	crying, grief, sad	136
Family	family	daughter, dad, aunt	118
Friends	friend	buddy, neighbor	95
Discrepancy	discrep	should, would	83
Feel	feel	feels, touch	128
Past focus	focuspast	ago, did, talked	341
Present focus	focuspresent	today, is, now	424
Future focus	focusfuture	may, will, soon	97
Time	time	end, until, season	310
Social processes	social	mate, talk, they	756
Achievement	achieve	win, success, better	213

The LIWC2015 word count was realised using PYTHON. The PYTHON script reads in .csv files containing one utterance per row. The specific categories mentioned in Table

2.1 were set and each row analyzed. Matching results incremented the LIWC2015 score and the output was written in a .csv file. In consequence of the LIWC2015 analysis at utterance-level, for each interview a .csv file was generated. These were combined in a post processing step based on patient language category and interventionist.

2.3 Data and Pre-Processing

2.3.1 Data

187 HDF5 (Hierarchical Data Format) files of transcribed motivational interviewing sessions from NIAAA were used. The HDF5 files were inspected with HDFView 3.0. HDF5 is a file format, data model and library for storing and managing data [59]. In general HDF5 files consist of datasets with raw data values itself and additional metadata which describes the data. Figure 2.3 shows the metadata of a HDF5 file out of the dataset.
The used HDF5 data contained information about the motivational interview session such as the words spoken, utterance time or the post-session client behavior. Figure 2.5 gives an example of the WORDS dataset from a used HDF5 file. The encoding of the HDF5 file replaces apostrophes and other signs with different symbols which were eliminated as a pre-processing step.

Figure 2.4 shows three datasets with their raw values. In the SPEAKER dataset (left) the current speaker can be identified, the CODE dataset (middle) shows the specific MISC code of the utterance and the WORDS dataset shows the spoken words. For example, at position 1 the patient (P) says "You're welcome" which is marked as follow neutral (FN).

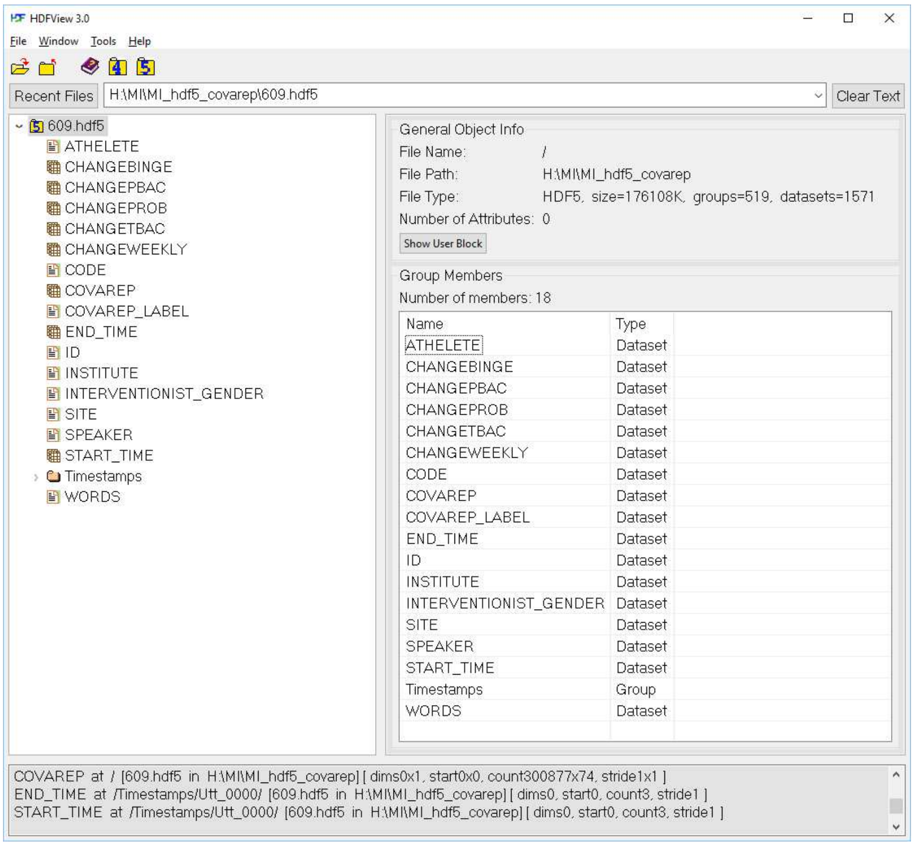

Figure 2.3: HDF5 metadata viewed with HDFView 3.0

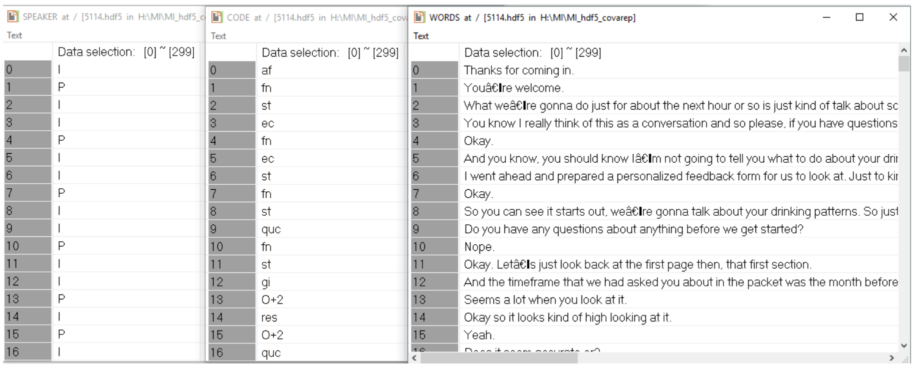

Figure 2.4: Dataset example of SPEAKER (left), CODE (middle) and WORDS (right)

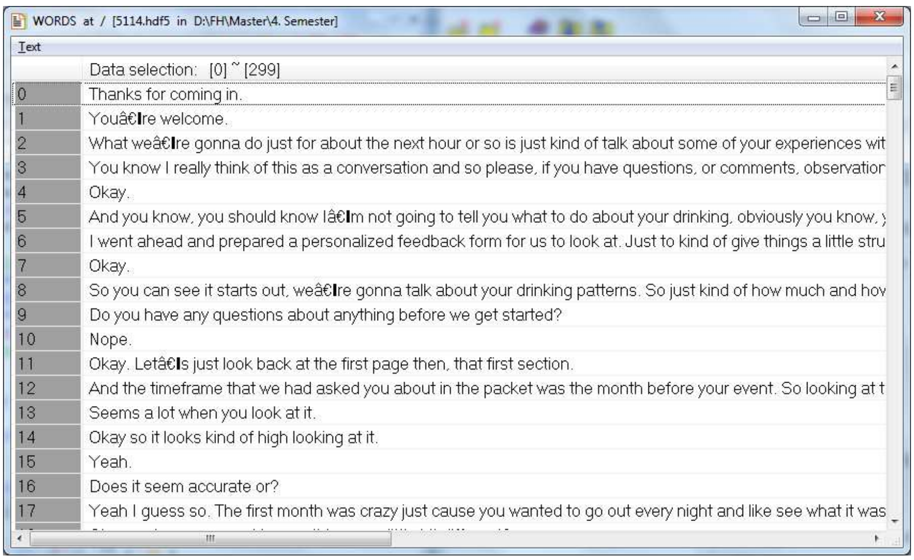

Figure 2.5: Raw data values of the WORDS dataset of Figure 2.3

The used and processed HDF5 data in this work were:

CHANGEBINGE A negative value if the patient improved the drinking behavior after the MI session. It is a positive value if the drinking related problems increased and zero if there was no change

CODE MISC code of the specific utterance

END_TIME Utterance end time

ID Patient ID

SPEAKER Interventionist (I) or patient (P)

START_TIME Utterance start time

WORDS Spoken words during the MI session

2.3.2 Pre-Processing

The main focus of the work was FN, CT and ST of the patient language as well as interventionist language, namely WORDS dataset from the HDF5 files. THE HDF5 data was read in with RStudio using the packages *hdf5r* and *rhdf5*.

To set the basis for follow-up tasks scripts were written to extract specific HDF5 data. Listing 2.1 shows an example of a function to read in the HDF5 data.
In order to obtain the utterances which are coded with e.g. FN the function shown in Listing 2.2 was used.

The functions which are described in Listing 2.1 to Listing 2.3 and modifications thereof, were used to gather data and utterances. In addition, the differentiation between Changer and Non-Changer was made. Subsequently each of the resulting utterances was cleared from misconceived characters occurring through encoding issues. Listing 2.3 shows the function to eliminate wrong characters using regular expressions. The utterances were saved in .txt files or .csv files respectively. Furthermore, methods were used to create files containing patient and interventionist utterances at interview-level and utterance-level. Files based at interview-level contained all utterances of interventionist/patient language from one MI-session additional categorized in FN, ST, CT in case of patient language. Files based at utterance-level contained only a single utterance per file.

Listing 2.1: R function getData

```

# Function to get HDF5 data

#file = HDF5 file name, data = e.g. "CODE" to get utterance codes
#person = "P" for patient, "I" for interventionist

getData <- function(file, data, person) {
  path <- "H:/MI/MI_hdf5_covarep/"
  datapath <- paste0(path, file)

  file = h5file(datapath, mode = "r")  # r = read only
  names = list.datasets(file) #get names from file e.g. "CODE"

  result = NULL

  #return data from patient and interventionist if no person is specified
    if (missing(person)) {
    for (i in 1:length(names)) {
      if (names[i] == data) {
        result <- h5read(datapath, data)
      }
    }
    return (result)
  } else{
    #return data from person
    speaker <- h5read(datapath, "SPEAKER")
    temp <- h5read(datapath, data)
    k = 1

    for (i in 1:length(speaker)) {
      if (speaker[i] == person) {
        result[k] <- temp[i]

        k = k + 1
      }
    }
    return(result)
  }
}
```

Listing 2.2: R function getFN

```

# Function to get follow neutrals

#person = "P" for patient, "I" for interventionist
getFnWords <- function(data, person) {
  path <- "H:/MI/MI_hdf5_covarep/"
  datapath <- paste0(path, data)
  code <- h5read(datapath, "CODE")
  speaker <- h5read(datapath, "SPEAKER")
  words <- h5read(datapath, "WORDS")

  k = 1
  wordsFN = NULL
  #return FN from patient and interventionist if no person is specified
  if (missing(person)) {
    for (i in 1:length(speaker)) {
      if ((code[i] == "FN") || (code[i] == "fn")) {
        wordsFN[k] = words[i]
        k = k + 1
      }
    }
    return (wordsFN)
  }
  #return FN utterances from person
  for (i in 1:length(speaker)) {
    if ((speaker[i] == person) && ((code[i] == "FN") ||
                                    (code[i] == "fn"))){
      wordsFN[k] = words[i]
      k = k + 1
    }
  }
  return (wordsFN)
}
```

Listing 2.3: R function cleanText

```

# Function to eliminate encoding characters

cleanText <- function(x) {
 x  <-
    gsub(" [ ](?=[ ])|[^-_,A-Za-z0-9,.,\\s]+",
         "’",
         x,
         ignore.case = TRUE,
         perl = TRUE)
}
```

Chapter 3

Results

All in all, 187 MI-interviews from binge drinking patients were investigated. The mean duration of the MI-sessions is 50.18 minutes. The therapy was successful with 60 patients, i.e. 60 patients improved their drinking behavior after the MI-interview.
The MI-session was unsuccessful with 127 patients, that means that 127 patients did not change their behavior after the MI-session, or the behavior got worse.

It was examined which topics are present in the MI-interviews. For this purpose topic models were created. The ten top terms were computed and the percentage of topics between Changer and Non-Changer was compared.
Furthermore, the patients and interventionists language was analyzed for their sentiments and language. For the sentiment analysis the patient language was subclassified into follow neutral (FN), change talk (CT) and sustain talk (ST). The amount of words, which fall into specific sentiment categories was calculated. To determine if there is a significant difference between the amount of sentiments from Changer and Non-Changer, the Wilcoxon rank-sum-test was used. The Wilcoxon rank-sum-test was chosen in all analyses, because the data and the results were not normally distributed. Boxplots present the sentiments with a significant difference between Changer and Non-Changer.

Table 3.1 shows number, mean durations and number of utterances of the investigated MI-interviews.

Table 3.1: Data overview; P = patient, I = interventionist

	Changer	Non-Changer
No. interviews	60	127
Mean duration	50.83 min	49.87 min
Mean duration per person	P: 15.25 min	P: 16.6 min
	I: 25.6 min	I: 24.53 min
Mean no. utterances	P: 200.35	P: 188.06
	I: 251.18	I: 237.00
Mean no. FN	119.18	109.16
Mean no. CT	53.51	51.83
Mean no. ST	44.65	42.31

3.1 Topic Modeling

Topic modeling determines the most common topics which occur in patient and interventionist language. The topic modeling was realised with RStudio using the packages *TM* [60], *Topicmodels* [61] and *LDA* [62]. The used topic modeling technique bases on [63].

The fundamental step was splitting, preprocessing and reading each single utterance of the HDF5-Files in a single .txt file. This ensured the topic modeling at utterance level in order to assign each utterance to a topic.
The created files were loaded into a corpus. To enable the term analysis, the files had to be formatted. The formatting pipeline consisted of transformation of the text to lower case, removing punctuation and removing numbers. After the formatting process, stopwords were removed. Stopwords are words which are ignored in analysis because they are very common and have no relevance for the document content. In this work the stopword list SMART from [60] in combination with own defined common text words was used.
Figure 3.1 shows the text processing pipeline to prepare the corpus text for the topic modeling process.

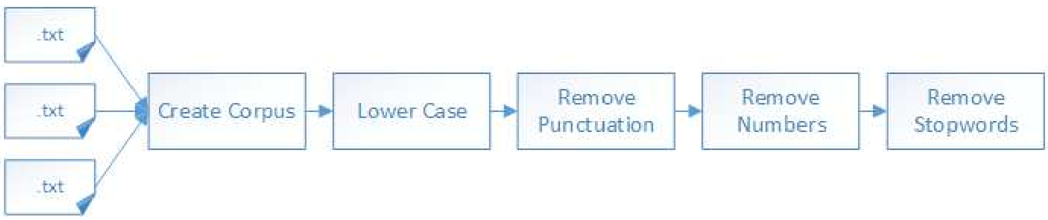

Figure 3.1: Pipeline to prepare the corpus text for topic modeling

Using the preprocessed corpus, a document-term matrix was created. Due to utterance level, removing stopwords resulted into null-entries of the corpus and therefore in the document-term matrix. These entries had to be removed respectively replaced by 0 before the topic modeling process.

The next step was executing Latent Dirichlet Allocation (LDA) using Gibbs sampling. Table 3.2 demonstrates the used LDA parameters.

The LDA output consisted of the top terms for each topic, the topic assignment for each document as well as the document probabilities being associated with each topic.

Table 3.3 shows the ten top terms of each topic.

Table 3.2: Used LDA Parameters

Parameter	Value	Description
burnin	0	Number of omitted Gibbs iterations at the beginning
iter	2000	Number of Gibbs iterations
thin	2000	Number of omitted in-between Gibbs iterations
nstart	5	Set random starts at 5
seed	254672 109 122887 145629037 2	Random integers as seed
best	true	Return the highest probability as the result
k	6	Number of calculated topics

Table 3.3: The ten top terms of the documents

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
drinking	drunk	night	feel	drink	people
beer	home	school	kind	time	friends
play	fine	week	good	drinks	party
fun	drive	times	pretty	alcohol	big
games	trouble	day	things	college	room
start	happen	high	bad	bit	kids
beers	stupid	drank	make	sick	campus
hard	person	remember	thought	hours	hang
stop	car	work	point	long	talk
game	problem	year	sense	average	stay

The terms listed in Table 3.3 could be summarized as following topics:

Topic 1 Drinking beer and playing games

Topic 2 Being drunk in combination with trouble regarding drink and drive

Topic 3 School and time reference

Topic 4 Feelings and emotions

Topic 5 Drinking alcohol at college

Topic 6 Party with friends and people

The topic model was created over the whole corpus which included all utterances from patient and interventionist language, except stopwords. Figure 3.2 displays the percentage of topics in interviews from Changer and from Non-Changer.

Figure 3.2 shows, that topic 1, topic 2 and topic 4 are the most common topics in interviews from Changer and Non-Changer. Non-Changer interviews are composed of 21.24 % topic 1 followed by 19.54 % topic 2 and 17.52 % topic 4. Changer interviews mainly consist of topic 2 with 19.54 % followed by topic 1 with 19.36 % and topic 4 with 17.52 %. However, the percentage of topics in the interviews are similar and there is no protruding topic.

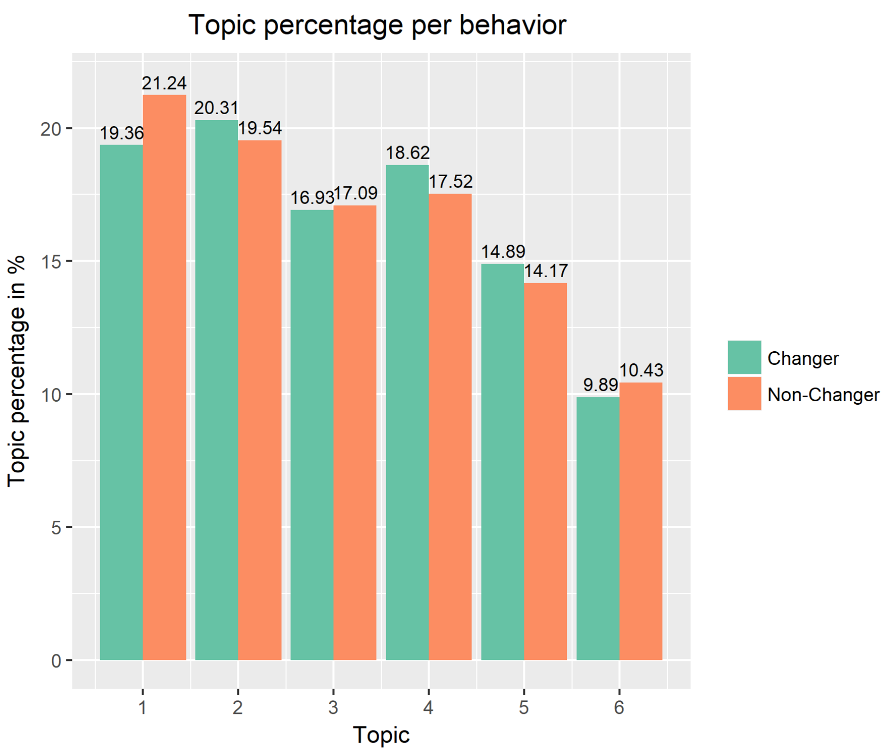

Figure 3.2: Percentage of topic amount in Changer and Non-Changer

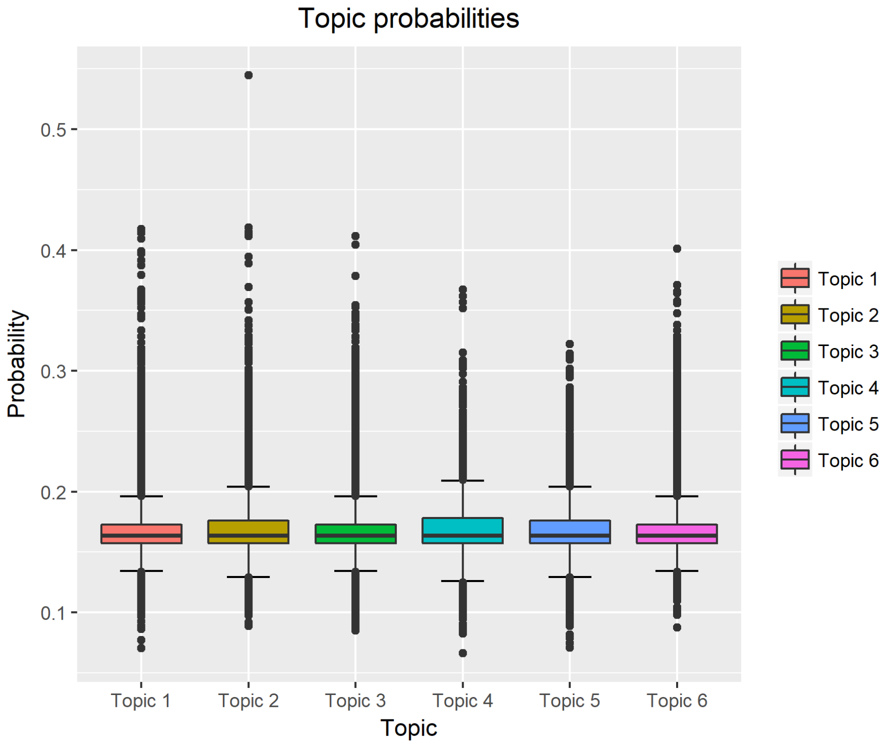

Figure 3.3: Boxplot of topic probabilities over all documents

Figure 3.3 shows the boxplots from the calculated topic probabilities over the corpus. Table 3.4 presents the detailed numbers. Topic 1 to topic 6 have the same median of 0.163 as well as the same first quartile with 0.157. Every topic has outliers, with topic 2 having the largest outlier with a maximum of 0.545.

Table 3.4: Topic probabilities summary

Topic	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Topic 1	0.070	0.157	0.163	0.166	0.173	0.418
Topic 2	0.089	0.157	0.163	0.167	0.176	0.545
Topic 3	0.085	0.157	0.163	0.167	0.173	0.412
Topic 4	0.066	0.157	0.163	0.167	0.178	0.367
Topic 5	0.071	0.157	0.163	0.167	0.176	0.322
Topic 6	0.088	0.157	0.163	0.166	0.173	0.401

3.2 Sentiment Analysis

The analysis of sentiments provides information about the emotions which occur in the interviews. It gives an understanding of the emotional intent of words and whether the utterance is positive, negative or characterized by other nuanced emotions. Two different methods were used to analyze the utterances. The first analysis was executed with the package *Tidytext* from R using the NRC library. The second method was text analyzing using the *LIWC* with PYTHON. LIWC allows to extend the text analysis and provides a broader insight into the used language.

Within both methods, the sentiments from Changer, Non-Changer and sentiments from the interventionist were calculated. Since the focus is on the patients language, sentiments were separately calculated for each of the MI-categories FN, ST ,CT and for the combination of all patient words. Calculation was based on interview level, that means the calculation was run for each MI-interview. To investigate, if there are differences in the emotions and used language between the parties, the Wilcoxon rank-sum test was applied. In order to provide comparability of the data between Changer and Non-Changer, the results of the sentiment analysis were normalized with the amount of utterances in the specific interviews.

The relationship between the sentiments from patient and interventionist was determined using Spearman's rho. Spearman's rho was chosen because the results of the sentiment analysis with NRC and LIWC had no normal distribution.

3.2.1 NRC Emotions

The NRC library investigates and assigns emotions to the words in one out of ten sentiments. The investigated sentiments are :

- anger
- anticipation
- negative
- positive
- disgust
- fear
- joy
- sadness
- surprise
- trust

The NRC sentiment analysis with *Tidytext* used .txt files as an input. For each MI-session a .txt file was created which contains the utterances of the MI-session. Additional subdivisions were conducted in dependency of patient language category FN, CT, ST and total patient language as well as total interventionist language in each case for Changer and Non-Changer. With help of *Tidytext* the files were tokenized. Tokenization creates a matrix where each utterance is divided into a single word. The resulting matrix contains each single word in a single row denoted with its utterance affiliation. With *anti_join()* stopwords were removed to decrease the number false positive counts (e.g. removing 'like' because it occurs frequently in conversations, as an example : "I used to measure them out with like a shot glass or whatever, I usually do it by eye, like I know, that's probably about like, two, when I have two. Like I, I do have like a generally good sense of what it is, but like once I start drinking, and if I try to make my own drink after that, I lose control"). After removing stopwords the sentiments were counted and presented in a dataframe. In addition to the sentiment counts, the number of utterances was added to the dataframe. The number of utterances is necessary to ensure a comparability of results because every interview has a different amount of utterances.

Table 3.5 shows the p-values of the Wilcoxon rank-sum-test for each sentiment of the patient language from Changer and Non-Changer. The significant level α is 0.05. Each value below α indicates a significant difference in the sentiment amount between Changer and Non-Changer. Significant differences have been found in the sentiments anger and negative in FN; anticipation, positive, fear, joy, surprise and trust in CT and in anger, anticipation, positive, disgust, joy and surprise in the total patient language, which represents all patient utterances regardless category. The sentiment analysis did not result in a significant p-value in any sentiment category from ST.

Table 3.5: p-values of the Wilcoxon rank-sum-test from patient language

	FN	CT	ST	total
anger	0.037	0.068	0.389	0.039
anticipation	0.109	0.032	0.152	0.034
negative	0.013	0.396	0.276	0.109
positive	0.121	0.004	0.065	0.021
disgust	0.095	0.070	0.213	0.023
fear	0.138	0.076	0.786	0.098
joy	0.062	0.003	0.057	0.005
sadness	0.332	0.333	0.128	0.154
surprise	0.184	0.002	0.349	0.009
trust	0.244	0.028	0.099	0.054

Table 3.6 presents the mean amount of each sentiment per utterance from Changer and Non-Changer in every patient language category. It is noticeable, that except for anger and fear in ST, the mean of Non-Changer is higher in every category than from Changer.

Table 3.6: Means of the number of sentiments per utterance from patient language; C = Changer, NC = Non-Changer

	FN		CT		ST		Total	
	C	NC	C	NC	C	NC	C	NC
anger	0.059	0.074	0.157	0.182	0.157	0.151	0.096	0.114
anticipation	0.167	0.196	0.304	0.367	0.371	0.423	0.232	0.272
negative	0.177	0.230	0.539	0.562	0.473	0.498	0.310	0.355
positive	0.283	0.321	0.397	0.489	0.505	0.610	0.342	0.404
disgust	0.05	0.064	0.144	0.175	0.130	0.146	0.084	0.103
fear	0.079	0.09	0.186	0.211	0.200	0.188	0.122	0.138
joy	0.156	0.185	0.239	0.304	0.344	0.413	0.202	0.247
sadness	0.083	0.097	0.226	0.241	0.193	0.217	0.133	0.152
surprise	0.064	0.073	0.114	0.152	0.127	0.141	0.085	0.104
trust	0.196	0.220	0.286	0.349	0.325	0.376	0.237	0.274

Table 3.7 shows the p-values of the wilcoxon rank-sum-test and the mean number of sentiments from the interventionist language. The significant level α from the wilcoxon rank-sum-test is 0.05. Each value below α indicates a significant difference in the amount of sentiments between the interventionist language in interviews from Changer and Non-Changer. Fear is the only sentiment with a significant p-value of 0.039.

Table 3.7: Wilcoxon rank-sum test p-value and means of the number of sentiments from interventionist language; C = interviews from Changer, NC = interviews from Non-Changer

	p-value	mean C	mean NC
anger	0.841	0.103	0.107
anticipation	0.807	0.397	0.394
negative	0.606	0.602	0.584
positive	0.548	0.755	0.761
disgust	0.958	0.105	0.106
fear	0.039	0.183	0.166
joy	0.064	0.349	0.375
sadness	0.663	0.173	0.169
surprise	0.808	0.167	0.172
trust	0.535	0.519	0.531

Figure 3.4 to Figure 3.6 show boxplots of the sentiment from patient language which have a significant p-value between the number of sentiments per utterance of Changer and Non-Changer.

Figure 3.4 represents the sentiments with a significant p-value of FN. On average, Non-Changer have more anger and negative sentiments per utterance than Changer.

Figure 3.5 presents the boxplot of CT sentiments. As in FN, CT has a higher number of significant sentiments per utterance than Non-Changer on average.

Figure 3.6 presents the significant sentiments over the total patient language. The highest mean difference can be seen in the positive sentiment. Changer has on average 0.342 less positive sentiments per utterance than Non-Changer with a mean value of 0.404.

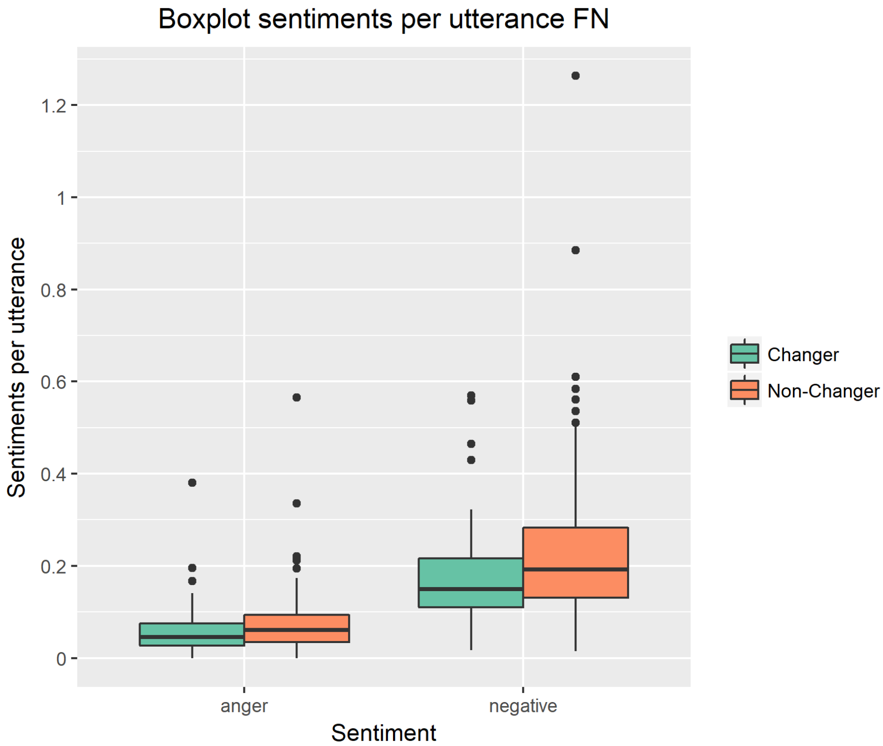

Figure 3.4: Boxplot of number of sentiments per utterance in FN with $p < 0.05$

Figure 3.7 shows the boxplot of the number of sentiments per utterance from interventionist language which have a significant median difference between the number of sentiments of Changer and Non-Changer. On average, interventionist from interviews with Changer, have 0.183 more fear related sentiments than interventionist in interviews with Non-Changer.

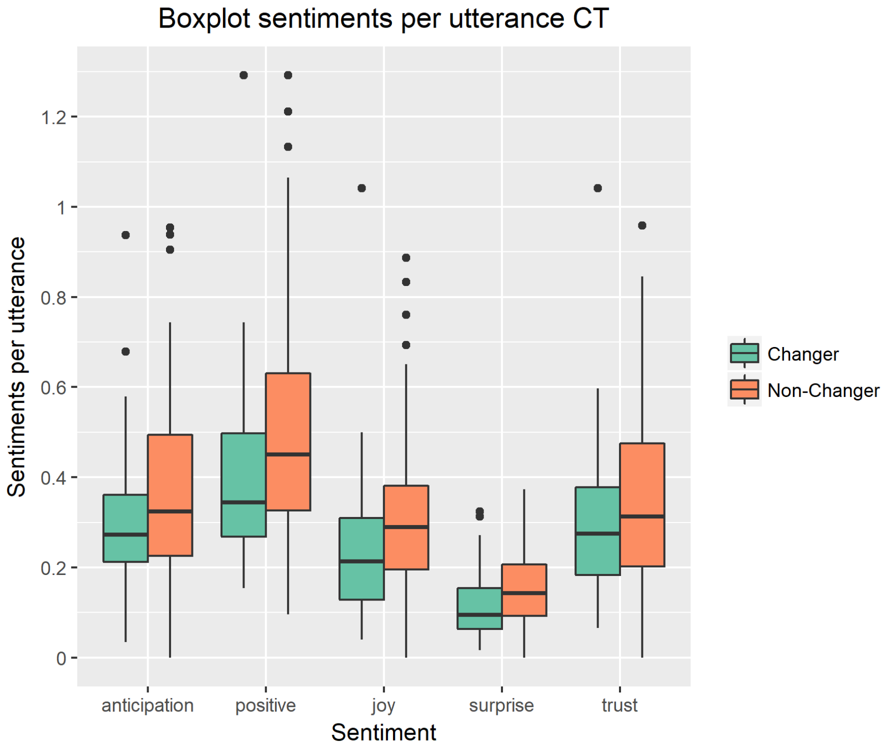

Figure 3.5: Boxplot of number of sentiments per utterance in CT with p <0.05

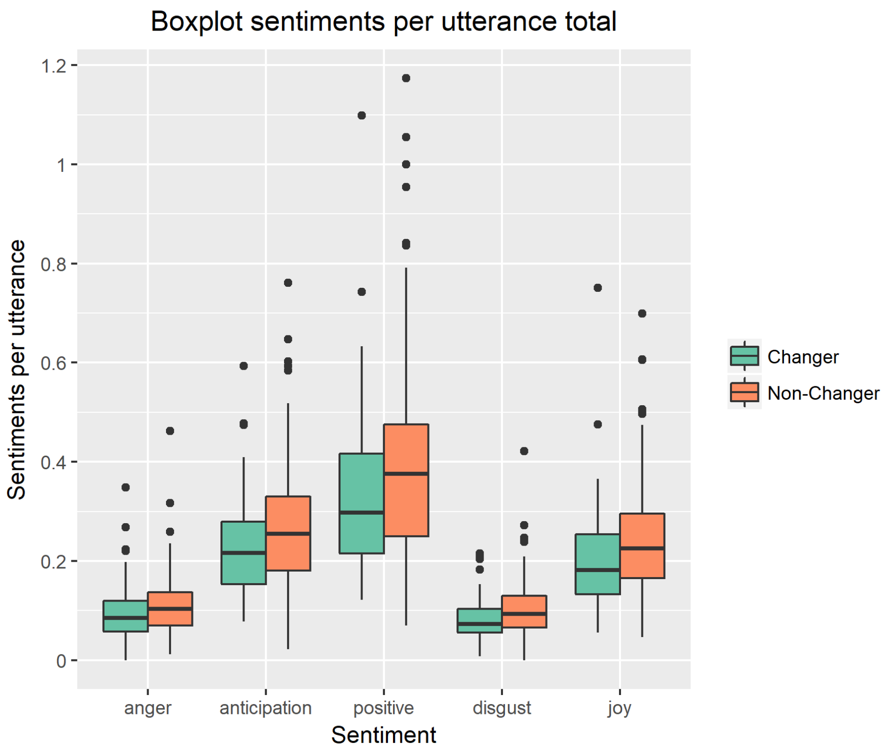

Figure 3.6: Boxplot of number of sentiments per utterance in total patient language with p <0.05

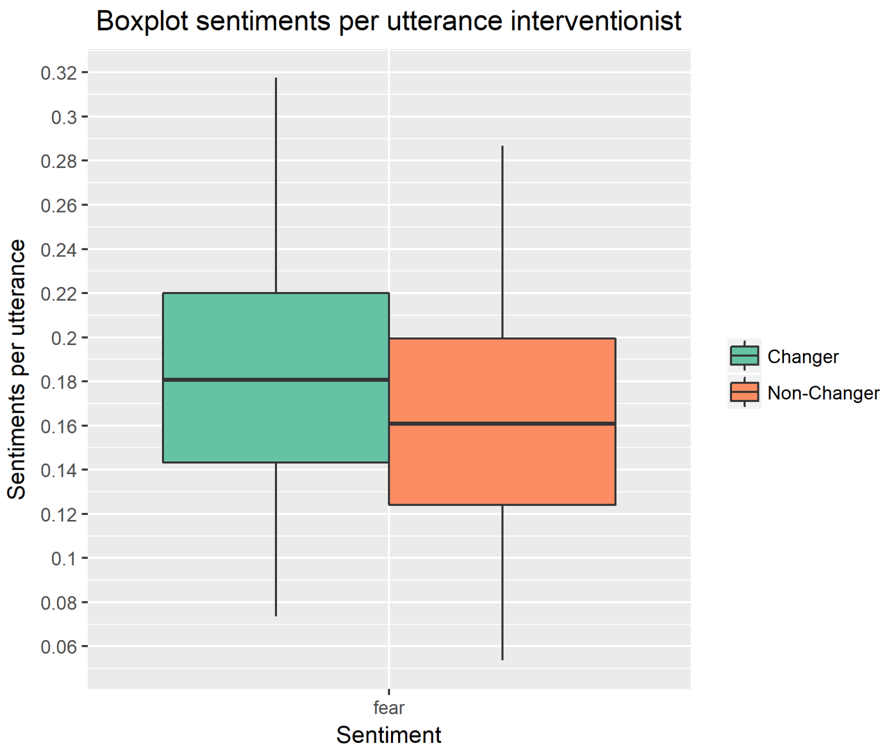

Figure 3.7: Boxplot of the number of sentiments per utterance of interventionist language with $p < 0.05$

Table 3.8 presents the Spearman's rho correlation coefficient between the sentiments from patient and interventionist. There is no correlation between the sentiments from patient and interventionist.

Table 3.8: Spearman's rho correlation between sentiments from patients and interventionists; C = interviews from Changer, NC = interviews from Non-Changer

	C	NC
anger	0.357	0.265
anticipation	0.260	0.165
disgust	0.356	0.185
fear	0.250	0.141
joy	0.368	0.301
negative	0.278	0.217
positive	0.403	0.267
sadness	0.478	0.162
surprise	0.337	0.116
trust	0.330	0.304

3.2.2 Linguistic Inquiry and Word Count

The LIWC library investigates and categorizes words in specific LIWC categories. The used categories and their abbreviation can be seen in Table 2.1.

The analysis with the LIWC library was realised using PYTHON. The python script uses .csv files which contains utterances. For each MI-session a .csv file was created which contains the utterances of the MI-session. As in the sentiment analysis with R, additional subdivisions were conducted in dependency of patient language category FN, CT, ST and total patient language as well as total interventionist language in each case for Changer and Non-Changer. Each row of the .csv file contains one utterance.
The python script enables the selection of specific LIWC categories and counts all words in the .csv file which belong to this category. The results were recorded in a new .csv file with the number of utterances and words.

For the Wilcoxon rank-sum test the LIWC counts were normalized at utterance level.

Table 3.9 shows the p-values of the Wilcoxon rank-sum test for each LIWC score of the patient language between Changer and Non-Changer. The significant level α is 0.05. Each value below α indicates a significant difference in the LIWC scores between Changer and Non-Changer. Significant differences have been found in the categories I, we, pronoun, sadness, focuspast, focuspresent, social and achievement in FN; positive and negative emotions as well as focuspast in CT and in we, negative emotions, focuspast and social in the total patient language, which presents all patient utterances regardless category. The LIWC scores did not result in a significant p-value in any LIWC category from ST.

Table 3.10 shows the mean LIWC scores per utterance of each LIWC category from Changer and Non-Changer in every category of the patient language.

Table 3.11 presents the resulting p-value from the Wilcoxon rank-sum test and the mean LIWC score per utterance from interventionist language from interviews with Changer and Non-Changer. In interventionist language the only category which resulted in a significant value was I with a p-value of 0.016.

Figure 3.8 represents the LIWC categories with a significant p-value of FN. On average, Non-Changer have a higher LIWC score per utterance in the categories I, we, pronoun, sadness, focus past, focus present, social and achievement than Changer.

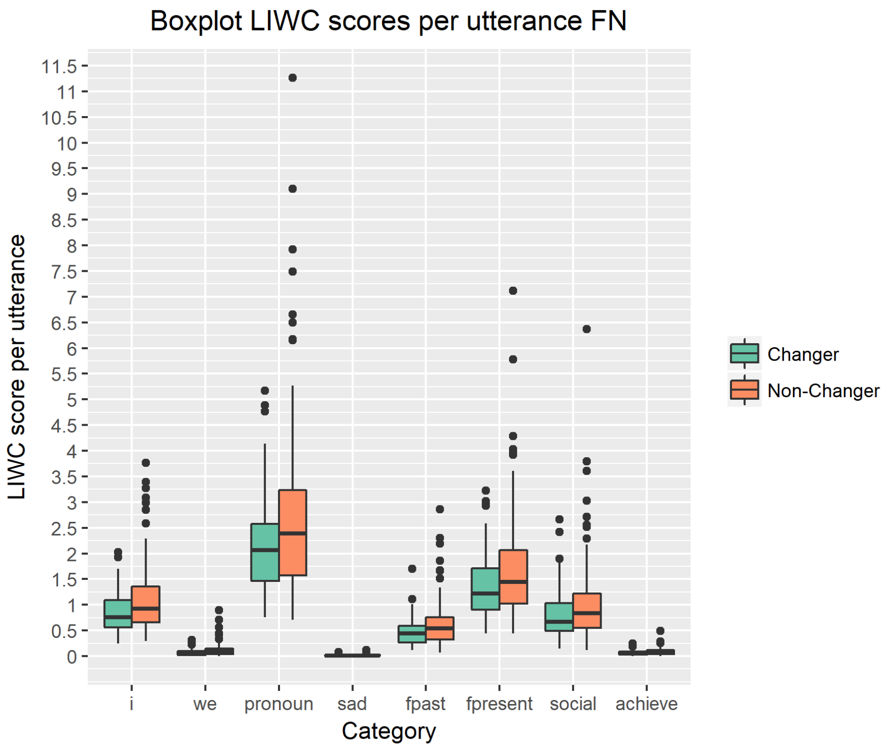

Figure 3.8: Boxplot LIWC scores per utterance of FN with p<0.05

Table 3.9: p-values of the Wilcoxon rank-sum test from patient language; FN = follow neutral, CT = change talk, ST = sustain talk, total = total patient language

	FN	CT	ST	total
i	0.029	0.360	0.150	0.070
we	0.013	0.055	0.750	0.033
shehe	0.321	0.464	0.593	0.333
they	0.118	0.405	0.527	0.103
pronoun	0.028	0.158	0.262	0.057
posemo	0.234	0.045	0.131	0.060
negemo	0.051	0.039	0.813	0.034
anx	0.244	0.495	0.078	0.785
anger	0.286	0.290	0.521	0.209
sad	0.013	0.427	0.374	0.151
family	0.750	0.890	0.812	0.733
friend	0.165	0.533	0.401	0.247
discrep	0.170	0.595	0.678	0.587
feel	0.194	0.976	0.704	0.186
focuspast	0.047	0.024	0.900	0.030
focuspresent	0.036	0.253	0.163	0.057
focusfuture	0.083	0.425	0.859	0.235
time	0.051	0.069	0.990	0.091
social	0.038	0.110	0.365	0.043
achieve	0.028	0.374	0.299	0.126

Table 3.10: Means of the LIWC scores from patient language; C = interviews from Changer, NC = interviews from Non-Changer

	FN		CT		ST		Total	
	C	NC	C	NC	C	NC	C	NC
i	0.873	1.094	1.840	1.957	1.934	2.207	1.278	1.492
we	0.078	0.122	0.068	0.114	0.090	0.087	0.078	0.114
shehe	0.044	0.065	0.054	0.072	0.034	0.016	0.045	0.059
they	0.087	0.117	0.102	0.130	0.098	0.118	0.093	0.119
pronoun	2.184	2.751	3.899	4.416	4.244	4.682	2.933	3.487
posemo	0.385	0.426	0.499	0.583	0.654	0.733	0.452	0.509
negemo	0.096	0.129	0.342	0.401	0.292	0.290	0.184	0.226
anx	0.011	0.014	0.051	0.056	0.062	0.040	0.029	0.030
anger	0.023	0.032	0.086	0.099	0.057	0.053	0.043	0.054
sad	0.013	0.020	0.050	0.055	0.037	0.043	0.026	0.033
family	0.015	0.017	0.025	0.028	0.013	0.016	0.017	0.019
friend	0.048	0.063	0.069	0.078	0.096	0.077	0.059	0.067
discrep	0.161	0.202	0.406	0.438	0.370	0.374	0.258	0.290
feel	0.052	0.058	0.134	0.132	0.170	0.172	0.092	0.094
focuspast	0.488	0.632	0.645	0.782	0.581	0.607	0.548	0.672
focuspresent	1.373	1.702	2.715	2.973	3.015	3.393	1.955	2.290
focusfuture	0.127	0.157	0.257	0.287	0.228	0.225	0.179	0.203
time	0.634	0.773	1.088	1.261	1.032	1.018	0.813	0.939
social	0.793	1.042	1.097	1.372	1.359	1.515	0.953	1.185
achieve	0.064	0.086	0.157	0.176	0.119	0.133	0.099	0.116

Table 3.11: P-values from Wilcoxon rank-sum test and means of the LIWC score per utterance from interventionist language; C = interviews from Changer, NC = interviews from Non-Changer

	p-value	mean C	mean NC
i	0.016	0.284	0.317
we	0.709	0.205	0.201
shehe	0.174	0.007	0.010
they	0.653	0.147	0.142
pronoun	0.172	4.599	4.780
posemo	0.310	0.854	0.881
negemo	0.576	0.229	0.223
anx	0.136	0.060	0.053
anger	0.445	0.023	0.025
sad	0.553	0.055	0.054
family	0.978	0.009	0.007
friend	0.089	0.042	0.050
discrep	0.906	0.333	0.334
feel	0.550	0.165	0.172
focuspast	0.407	0.758	0.787
focuspresent	0.120	2.751	2.869
focusfuture	0.967	0.361	0.359
time	0.901	1.089	1.077
social	0.136	2.804	2.949
achieve	0.756	0.177	0.178

In Figure 3.9 the LIWC categories with a significant p-value of CT are represented. On average, Non-Changer have a higher LIWC score per utterance in the categories of positive and negative emotions and focus past. The boxplot shows, that there is a higher LIWC score per utterance in positive emotions than in negative emotions on average. Non-Changer hold a higher score than Changer, however Changer have a maximum in positive emotions of 1.51 and Non-Changer 1.37 With a significant mean of 0.645 for Changer and 0.782 for Non-Changer, the LIWC category focus past has the highest LIWC score per utterance in CT.

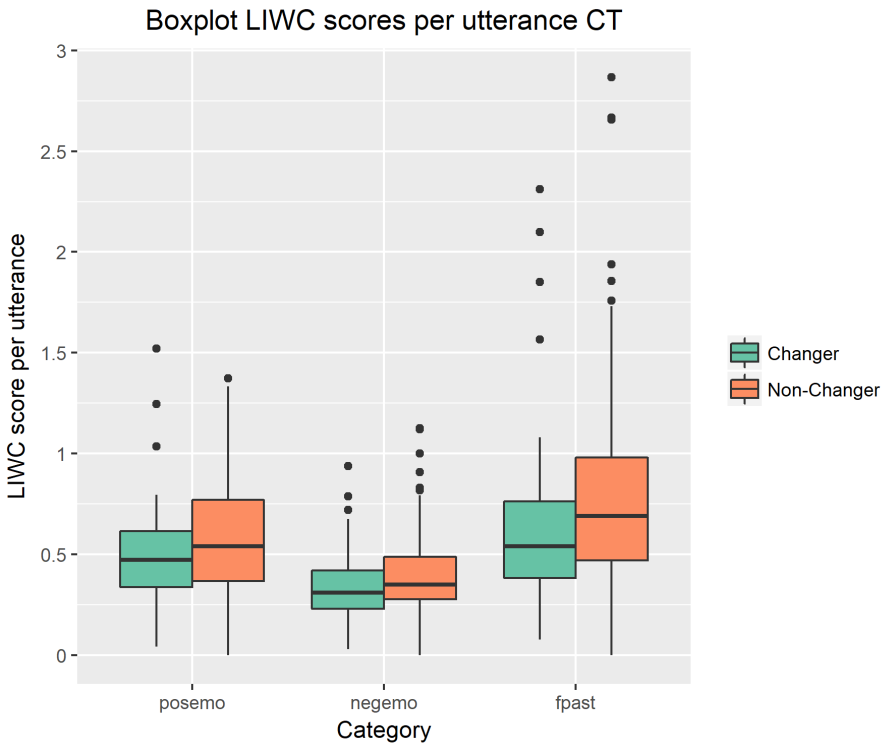

Figure 3.9: Boxplot LIWC scores per utterance of CT with p<0.05

Figure 3.10 gives information about the significant LIWC categories and their LIWC scores per utterances of the total patient language. The social related LIWC category demonstrates the highest mean with a 0.953 for Changer and 1.185 for Non-Changer.

The boxplot of the significant LIWC categories from interventionist language is presented in Figure 3.11. The only category with a significant p-value is I. Changer hold a mean of 0.284 and Non-Changer 0.317, however, Changer have higher outliers with a maximum of 0.897 and Non-Changer 0.797.

Table 3.12 shows the Spearman's rho correlation coefficient between the LIWC scores from patient and interventionist. There is no correlation between the LIWC scores from patient and interventionist.

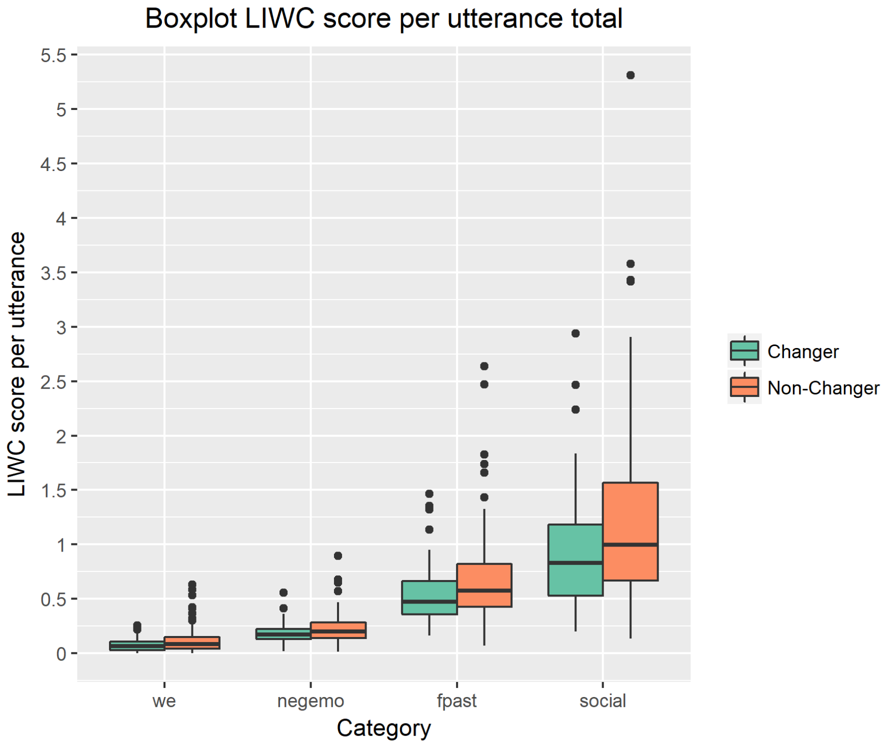

Figure 3.10: Boxplot LIWC scores per utterance of total patient language with p<0.05

Figure 3.11: Boxplot LIWC scores per utterance of total interventionist language with p<0.05

Table 3.12: Spearman's rho correlation between LIWC scores from patients and interventionists; C = interviews from Changer, NC = interviews from Non-Changer

	C	NC
i	0.511	0.171
we	0.064	0.01
shehe	0.587	0.456
they	0.229	0.089
pronoun	0.420	0.206
posemo	0.495	0.325
negemo	0.290	0.247
anx	0.308	0.292
anger	0.274	0.311
sad	0.343	0.238
family	0.546	0.567
friend	0.400	0.354
discrep	0.333	0.178
feel	0.419	0.271
focuspast	0.224	0.211
focuspresent	0.468	0.105
focusfuture	0.176	0.110
time	0.291	0.092
social	0.403	0.114
achieve	0.550	0.256

Chapter 4

Discussion

In this work transcribed motivational interview data from binge drinking patients were analysed. The patient language as well as the interventionist language were investigated with topic modeling and sentiment analysis.

4.1 Topic Modeling

The results in Table 3.3 show that Non-Changer have the most utterances related to topic 1 which makes up 21.24 %. Therefore, Non-Changer talk more about drinking in combination with games. The top terms in topic 1 imply that Non-Changer have fun drinking and playing drinking games and that they find it hard to stop. So one reason that the MI-session was not successful with this patient could be that Non-Changer have too much fun with their binge drinking behavior and thus they do not really want a change. Topic 1 is followed by topic 2 with 19.54 %. It includes subjects about drinking related with problems and trouble especially driving drunk.

Topic 1 and topic 2 are also the most leading topics in Changer. However, Changer have higher percentage of 20.31 % in topic 2 followed by 19.36 % in topic 1. This indicates that Changer talk more about the problems that occur with drinking and not so much about the positive aspects like fun.

Topic 4 follows topic 1 and topic 2 in both Changer and Non-Changer. Changer interviews contain 18.62 % topic 4 and Non-Changer 17.52 %. That shows that Changer talk more about their feelings than Non-Changer.

Topic 3, topic 5 and topic 6 do not show sticking out differences between Changer and Non-Changer.

The topic model results show that Changer talk more about drinking related problems and focus on positive feelings. However, the percentage differences between Changer and Non-Changer are very low and therefore no clear statement can be made. Moreover, it must be taken into account that stopwords were removed. *No* or *not* are marked as stopwords and therefore utterances containing for example "not drinking" were not considered.

4.2 Sentiment Analysis

Sentiment analysis were executed using the two different libraries NRC and LIWC.

4.2.1 NRC

The patient language was analysed by subdividing the coded utterances into FN, ST and CT. The Wilcoxon rank-sum test was used with a significance level of $\alpha = 0.05$ and with normalized sentiments on utterance level.

The Wilcoxon rank-sum test of the FN category showed, that there are significant differences in the investigated sentiments in anger and in negative sentiments. The mean of anger in Non-Changer was higher ($\mu = 0.074$) than in Changer ($\mu = 0.059$). In addition, the mean of Non-Changer in negative sentiments was higher ($\mu = 0.230$) than from Changer ($\mu = 0.177$). Thus, Non-Changer have more anger and negative utterances than Changer on average. These results are in contradiction with results from Project MATCH (cf. [64]), which claims that Changer are angrier than Non-Changer.

Utterances from the category of CT showed significant differences between Changer and Non-Changer in anticipation, positive, fear, joy, surprise and trust. In each of this sentiment category, the mean of Non-Changer is higher than the mean of Changer. That indicates that Non-Changer use on average more words of anticipation, positive, joy, surprise and trust. This raises questions, because one might assume that Changer would use more positive language than Non-Changer.

Sentiment analysis in ST showed no significant differences.

Investigating the total patient language, the Wilcoxon rank-sum test showed a significant sentiment difference in Non-Changer and Changer in the categories anger, anticipation, positive, disgust, joy and surprise. That shows, that beside the negative sentiments anger and disgust even the positive sentiments like anticipation, joy and surprise play a role. The fact that in every significant category Non-Changer have a higher mean than Changer could be due to the fact that Non-Changer talk more and therefore use more sentiment words which can be analyzed during the interviews than Changer (mean talking duration Non-Changer patients 16.60 minutes vs. 15.25 minutes Changer patients).

The interventionist language was analysed without subdividing into different categories. The Wilcoxon rank-sum test showed a significant difference of interventionist language between Changer and Non-Changer in the category of fear. Interventionists used more fear related language in interviews with Changer ($\mu = 0.183$) than in interviews with Non-Changer ($\mu = 0.166$). This result is unexpected because one might interpret using fear related sentiments as an motivational interviewing inconsistent behavior. It is also possible that the interventionist might reflect patient language, even when Wilcoxon rank-sum test do not show significant values in fear among patient language, leading in a higher occurrence of sentiments per utterances in Changer interviews than in Non-Changer interviews.

It should also be considered that stopwords were removed. One stopword which was removed was "like". This word was removed because it occurred frequently as a colloquial word during interviews and would have led to a huge amount of false sentiment calculations in positive categories. Furthermore, the NRC is a unigram based method and does not take qualifiers into account. This means that utterances like "not good" or "not bad" have not been considered.

The Spearmans's rho correlation between the NRC sentiments from patient and interventionists shows a low correlation in all categories from Changer and Non-Changer. It is noticeable though, that on average the correlation coefficient between patient and interventionist is higher in Changer than in Non-Changer.

4.2.2 LIWC

The patient language was analysed by subdividing the coded utterances into FN, ST and CT. The Wilcoxon rank-sum test was used with an significance level of $\alpha = 0.05$ and with normalized sentiments on utterance level.

The Wilcoxon rank-sum test of the FN category showed, that there are significant differences in the investigated LIWC counts in I, we, pronoun, negative emotions, sadness, focus past and focus present, social and achievement. As with NRC, the mean of LIWC counts of Non-Changer is higher in every significant category than the mean of Changer. The higher LIWC scores of pronouns indicate that Non-Changer use significant more pronouns than Changer (category pronoun mean Non-Changer $\mu = 2.75$ vs. mean Changer $\mu = 2.18$).

The Wilcoxon rank-sum test used on the LIWC counts from CT showed p-values smaller than 0.05 in the categories negative emotions, positive emotions and focus past. The utterances from Non-Changer contain on average more negative emotions ($\mu = 0.40$) than the utterances from Changer ($\mu = 0.342$). Furthermore, Non-Changer talk more ($\mu = 0.782$) in the past than Changer ($\mu = 0.645$).

The category ST did not show any significant p-values using Wilcoxon rank-sum test. This discovery matches with missing significance using the NRC method.

In order to analyse pronouns and time related categories, no stopwords were removed.

Investigating the total patient language, the Wilcoxon rank-sum test showed a significant LIWC score difference in Non-Changer and Changer in the categories we, negative emotion, focus past and social. Including the total patient language, Non-Changer show a mean of $\mu = 0.226$ in negative emotions while Changer hold a mean of $\mu = 0.184$, that indicates that Non-Changer are more negative than Changer.

As with the NRC method, the interventionist language was analysed without subdividing it into different categories. The Wilcoxon rank-sum test resulted in a p-value of 0.016 in the category of I with a mean in Changer interviews of $\mu = 0.284$ and a mean in Non-Changer interviews of $\mu = 0.317$. According to this, interventionists use more I pronouns in interviews with Changer than in interviews with Non-Changer.

The Spearmans's rho correlation between LIWC scores form patient and interventionists shows a medium correlation in the categories I, she/he and family from Changer, as well as in family of Non-Changer. The higher correlation in the category of I in Changer ($\rho = 0.511$) than in Non-Changer ($\rho = 0.171$) adumbrate a medium correlation between patient and interventionist. All other categories show low positive correlation.

Both methods did not show any significant results in the patient language category ST. One might assume that in ST negative emotions outweigh, but that was not the case. One possible reason that ST did not show any significants could be coding issues within ST.

The Wilcoxon rank-sum test showed in both methods significant values in the category negative emotions, although they were not in the same patient language category (NRC in FN, LIWC in CT).

Chapter 5

Conclusion

In this work 187 motivational interview sessions from binge drinking patients were analysed and patient language as well as the interventionist language investigated. The language was examined with topic models and with sentiment analysis with the goal, to find remarkable language differences between people who change their behavior after the MI-session and people who do not change. It was assumed that the results would lead to a small part on answering the question why motivational interviewing works.

The topic modeling showed that people with a positive post-session behavior change, talk more about their problems and troubles which can be attributed to drinking. They also talk more about emotions. People for whom the therapy was not successful had the positive aspects of drinking related to fun as a main topic. However, the percentage differences between the topics were very low and therefore no clear statement can be made.

Both sentiment analysis methods resulted in significant differences in patient language between people with and without post-session behavior change. Although the methods are not comparable, they showed significant differences in negative sentiments categories like anger, disgust or sadness and related. Especially Non-Changer resulted in a higher amount of those sentiments per utterance which contradicts findings from another study.

Follow neutral is still a very neglected category when it comes to analysis of MI patient language. A not negligible amount of significant results can be seen in the patient category follow neutral. These results could suggest, that follow neutral is a key factor in understanding motivational interviewing and more research in follow neutral direction is needed.

Despite the widespread usage of motivational interviewing its underlying mechanism is still poorly understood. Further research investigation in patient language especially follow neutral could lead to a better understanding of the mechanism of motivational interviewing. Investigating interventionist language especially the reciprocal influence of patient and interventionist language, could allow a more precise understanding of MI mechanism and could lead to a better prediction of post-session behavior change.

A next investigation step could be the analysis of the sound of the non-verbal acoustic and dyadic speech indicators. This investigation would go beyond what is said and take into account how and in what interpersonal context something is said. Such an approach could allow to code sessions almost in real time and therefore provide more accurate predictions about post-session behavior change.

List of Abbreviations

ALT	Alanine Amino Transferase
AST	Aspartate Amino Transferase
AUDIT	Alcohol Use Disorders Identification Test
AUQ	Alcohol Use Questionnaire
BAC	Blood Alcohol Concentration
BOW	Bag of Words
BRFSS	Behavioral Risk Factor Surveillance System
C	Changer
CAS	Harvard School of Public Health College Alcohol Study
CDC	Centers for Disease Control and Prevention
CDT	Carbohydratedeficient Transferrin
CHARM	Contextual Hierarchical Attention-based Recurrent Model
CRISP-DM	Cross Industry Standard Process for Data Mining
CT	Change Talk
DSF	Discrete Sentence Feature
EM	Expectation-Maximization
EtG	Ethyl Glucuronide
EtS	Ethyl Sulfate
FN	Follow Neutral
GGT	Gamma Glutamyl Transferase
HDF5	Hierarchical Data Format
HIV	Human Immunodeficiency Virus
IE	Information Extraction
IR	Information Retrieval
LDA	Latent Dirichlet Allocation

LIWC	Linguistic Inquiry and Word Count
LSA	Latent Semantic Analysis
LSI	Latent Semantic Indexing
MCV	Mean Corpuscular Volume
MI	Motivational Interviewing
MICO	Motivational Interview Consistent
MIIN	Motivational Interview Inconsistent
MISC	Motivational Interviewing Skill Code
MITI	Motivational Interviewing Treatment Integrity
NC	Non-Changer
NIAAA	National Institute on Alcohol Abuse and Alcoholism
NLP	Natural Language Processing
NRC	National Research Council Canada
NSDUH	National Survey on Drug Use and Health
PEth	Phosphatidylethanol
PLSA	Probabilistic Latent Semantic Analysis
PLSI	Probabilistic Latent Semantic Indexing
RNN	Recursive Neural Network
SAMHSA	Substance Abuse and Mental Health Services Administration
SCOPE	Sequential Code for Observing Process Exchanges
ST	Sustain Talk
SVD	Singular-Value Decomposition
TBC	Target Behavior Change
TLFB	Timeline Followback
WHO	World Health Organization

List of Figures

List of Tables

Bibliography

[1] N. I. on Alcohol Abuse and A. (NIAAA), "Alcohol facts and statistics." https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/alcohol-facts-and-statistics. Accessed:2018-04-10. 1, 1.1.2

[2] K. M. Jennison, "The short-term effects and unintended long-term consequences of binge drinking in college: a 10-year follow-up study," *The American journal of drug and alcohol abuse*, vol. 30, no. 3, pp. 659–684, 2004. 1, 1.1.2

[3] N. I. on Alcohol Abuse and A. (NIAAA), "Drinking levels defined." https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking, 2012. Accessed:2018-04-10. 1, 1.1.1

[4] H. Wechsler and T. F. Nelson, "Binge drinking and the american college students: What's five drinks?," *Psychology of Addictive Behaviors*, vol. 15, no. 4, p. 287, 2001. 1.1

[5] M. R. Piano, A. Mazzuco, M. kang, and S. A. Phillips, "Binge drinking episodes in young adults: How should we measure them in a research setting?," *Journal of Studies on Alcohol and Drugs*, vol. 78, no. 4, pp. 502–511, 2017. 1.1, 1.1.1, 1.1.1, 1.1

[6] N. I. on Alcohol Abuse and A. (NIAAA), "College drinking." https://pubs.niaaa.nih.gov/publications/CollegeFactSheet/Collegefactsheet.pdf. 1.1.1

[7] W. H. Organization, "Standard drink measures, in grams per unit." http://apps.who.int/gho/data/view.main.54180. Accessed:2018-07-09. 1.1.1

[8] E. Roehm, "Nutrition heart - alcoholic drinks and grams of alcohol." http://www.nutritionheart.com/alcohol-drinks-grams-of-alcohol. Accessed:2018-07-09. 1.1.1

[9] W. H. Organization, "Heavy episodic drinking among drinkers - situation and trends." http://www.who.int/gho/alcohol/consumption_patterns/heavy_episodic_drinkers_text/en/. Accessed:2018-06-27. 1.1.1

[10] B. R. F. S. System. https://health.mo.gov/data/brfss/definition.php. Accessed:2018-07-09. 1.1.1

[11] S. Abuse and M. H. S. Administration, "Results from the 2013 national survey on drug use and health: Summary of national findings," 2014. 1.1.1

[12] W. H. Organization, "Audit: The alcohol use disorders identification test." http://www.who.int/substance_abuse/activities/sbi/en/. Accessed:2018-07-02. 1.1.1

[13] B. et al., "Audit," *World Health Organization*. 1.1.1

[14] J. Townshend and T. Duka, "Patterns of alcohol drinking in a population of young social drinkers: a comparison of questionnaire and diary measures," *Alcohol and Alcoholism*, vol. 37, no. 2, pp. 187–192, 2002. 1.1.1

[15] S. Abuse, M. H. S. Association, *et al.*, "The role of biomarkers in the treatment of alcohol use disorders, 2012 revision," *Advisory*, vol. 11, no. 2, pp. 1–8, 2012. 1.1.1

[16] H. Wechsler, G. Dowdall, and M. Moeykens, "Health and behavioral consequences of binge drinking in college," *Jama*, vol. 272, pp. 1672–1677, 1994. 1.1.2

[17] E. Kuntsche, S. Kuntsche, J. Thrul, and G. Gmel, "Binge drinking: health impact, prevalence, correlates and interventions," *Psychology & health*, vol. 32, no. 8, pp. 976–1017, 2017. 1.1.2

[18] H. Wechsler and T. F. Nelson, "What we have learned from the harvard school of public health college alcohol study: Focusing attention on college student alcohol consumption and the environmental conditions that promote it," *Journal of Studies on alcohol and Drugs*, vol. 69, no. 4, pp. 481–490, 2008. 1.1.2, 1.1.3

[19] G. W. Dowdall and H. Wechsler, "Studying college alcohol use: widening the lens, sharpening the focus.," *Journal of Studies on Alcohol, Supplement*, no. 14, pp. 14–22, 2002. 1.1.3, 1.1

[20] M. Tanana, K. A. Hallgren, Z. E. Imel, D. C. Atkins, and V. Srikumar, "A comparison of natural language processing methods for automated coding of motivational interviewing," *Journal of substance abuse treatment*, vol. 65, pp. 43–50, 2016. 1.2, 1.2.1, 1.2.3

[21] J. Hettema, J. Steele, and W. R. Miller, "Motivational interviewing," *Annu. Rev. Clin. Psychol.*, vol. 1, pp. 91–111, 2005. 1.2, 1.2.1

[22] W. R. Miller, "Motivational interviewing: research, practice, and puzzles," *Addictive behaviors*, vol. 21, no. 6, pp. 835–842, 1996. 1.2.1

[23] W. R. Miller and S. Rollnick, "Ten things that motivational interviewing is not," *Behavioral and cognitive psychotherapy*, vol. 37, no. 2, pp. 129–140, 2009. 1.2.1

[24] W. R. Miller and S. Rollnick, *Motivational interviewing: Helping people change*. Guilford press, 2012. 1.2.1, 1.2.2

[25] L. H. Glynn and T. B. Moyers, "Chasing change talk: The clinician's role in evoking client language about change," *Journal of substance abuse treatment*, vol. 39, no. 1, pp. 65–70, 2010. 1.2.1

[26] M.-C. C. J. W. S. S. B. B. Xiaolei Huang, Lixing Liu, "Learning dyadic attention networks to predict outcomes of motivational interviewing," 1.2.1, 1.2.3

[27] S. Rollnick and W. R. Miller, "What is motivational interviewing?," *Behavioral and Cognitive Psychotherapy*, vol. 23, no. 4, p. 325–334, 1995. 1.2.2

[28] T. Martin, T. B. Moyers, J. Houck, P. Christopher, and W. R. Miller, "Motivational interviewing sequential code for observing process exchanges (mi-scope) coder's manual," *Retrieved March*, vol. 16, p. 2009, 2005. 1.2.3, 1.2.3, 1.2.3, 1.2.3

[29] S. A. Center on Alcoholism and Addictions, "Motivational interviewing skills code (misc) 2.1." https://casaa.unm.edu/code/misc.html. Accessed:2018-07-25. 1.2.3

[30] W. Miller, T. Moyers, D. Ernst, and P. Amrhein, "Manual for the motivational interviewing skill code (misc) version 2.1, 2008," *University of New Mexico Center on Alcoholism Substance Abuse, and Addictions website: https://casaa. unm. edu/codinginst. html. Accessed*, vol. 20, 2017. 1.2.3, 1.2.3, 1.2.3

[31] S. A. Center on Alcoholism and Addictions, "Sequential code for observing process exchanges (scope)." https://casaa.unm.edu/codinginst.html. Accessed:2018-07-26. 1.2.3

[32] T. B. Moyers and T. Martin, "Therapist influence on client language during motivational interviewing sessions," *Journal of substance abuse treatment*, vol. 30, no. 3, pp. 245–251, 2006. 1.2.3, 1.2.3

[33] T. B. Moyers, T. Martin, J. K. Manuel, S. M. Hendrickson, and W. R. Miller, "Assessing competence in the use of motivational interviewing," *Journal of substance abuse treatment*, vol. 28, no. 1, pp. 19–26, 2005. 1.2.3

[34] T. Moyers, J. Manuel, and D. Ernst, "Motivational interviewing treatment integrity coding manual 4.1," *Unpublished manual*, 2014. 1.2.3

[35] S. A. Center on Alcoholism and Addictions, "Motivational interviewing treatment integrity (miti) code 4.2." https://casaa.unm.edu/codinginst.html. Accessed:2018-07-26. 1.2.3

[36] R. Plutchik, "The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," *American scientist*, vol. 89, no. 4, pp. 344–350, 2001. 1.2, 1.3, 1.3, 1.4

[37] A. Ben-Zeev, "The nature of emotions," *Philosophical Studies*, vol. 52, no. 3, pp. 393–409, 1987. 1.3

[38] M. Allahyari, S. Pouriyeh, M. Assefi, S. Safaei, E. D. Trippe, J. B. Gutierrez, and K. Kochut, "A brief survey of text mining: Classification, clustering and extraction techniques," *arXiv preprint arXiv:1707.02919*, 2017. 2.1, 2.1.1

[39] S. V. Europe, "What is the crisp-dm methodology." https://www.sv-europe.com/crisp-dm-methodology. Accessed:2018-04-10. 2.1

[40] R. Alghamdi and K. Alfalqi, "A survey of topic modeling in text mining," *Int. J. Adv. Comput. Sci. Appl.(IJACSA)*, vol. 6, no. 1, 2015. 2.1.2

[41] J. Chang, S. Gerrish, C. Wang, J. L. Boyd-Graber, and D. M. Blei, "Reading tea leaves: How humans interpret topic models," in *Advances in neural information processing systems*, pp. 288–296, 2009. 2.1.2

[42] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003. 2.1.2, 2.1.2, 2.1.2, 2.1.2, 2.1.2, 2.2

[43] S. Huh and S. E. Fienberg, "Discriminative topic modeling based on manifold learning," *ACM Trans. Knowl. Discov. Data*, vol. 5, pp. 20:1–20:25, feb 2012. 2.1.2

[44] T. Hofmann, "Unsupervised learning by probabilistic latent semantic analysis," *Machine learning*, vol. 42, no. 1-2, pp. 177–196, 2001. 2.1.2

[45] T. K. Landauer, P. W. Foltz, and D. Laham, "An introduction to latent semantic analysis," *Discourse processes*, vol. 25, no. 2-3, pp. 259–284, 1998. 2.1.2

[46] P. Wiemer-Hastings, K. Wiemer-Hastings, and A. Graesser, "Latent semantic analysis," in *Proceedings of the 16th international joint conference on Artificial intelligence*, pp. 1–14, Citeseer, 2004. 2.1.2

[47] W. M. A. M. Zubir, I. A. Aziz, and J. Jaafar, "A survey on textual semantic classification algorithms," in *2017 IEEE Conference on Big Data and Analytics (ICBDA)*, 2017. 2.1, 2.1.2, 2.1.2, 2.1.2, 2.1.2, 2.2

[48] T. Brants, F. Chen, and I. Tsochantaridis, "Topic-based document segmentation with probabilistic latent semantic analysis," in *Proceedings of the Eleventh International Conference on Information and Knowledge Management*, CIKM '02, (New York, NY, USA), pp. 211–218, ACM, 2002. 2.1.2, 2.1.2

[49] H. Jelodar, Y. Wang, C. Yuan, and X. Feng, "Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey," *arXiv preprint arXiv:1711.04305*, 2017. 2.1.2

[50] B. Liu, "Sentiment analysis and opinion mining: Synthesis lectures on human language technologies.[si]: Morgan & claypool publishers, 2012. 168 p," 2016. 2.2

[51] S. M. Mohammad and P. D. Turney, "Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon," in *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*, pp. 26–34, Association for Computational Linguistics, 2010. 2.2

[52] H. Wickham and L. Henry, *tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions*, 2018. R package version 0.8.0. 2.2.1

[53] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word-emotion association lexicon," vol. 29, no. 3, pp. 436–465, 2013. 2.2.1

[54] Hu and Liu, "Opinion lexicon," 2004. 2.2.1

[55] F. Å. Nielsen, "Afinn," mar 2011. 2.2.1

[56] J. Silge and D. Robinson, "Text mining with r," 2.2.1

[57] H. Wickham, R. Francois, L. Henry, and K. Müller, *dplyr: A Grammar of Data Manipulation*, 2017. R package version 0.7.4. 2.2.1

[58] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, "The development and psychometric properties of liwc2015," tech. rep., 2015. 2.2.2

[59] T. H. Group, "Hdf5." https://portal.hdfgroup.org/display/HDF5/HDF5. Accessed:2018-05-03. 2.3.1

[60] I. Feinerer and K. Hornik, *tm: Text Mining Package*, 2017. R package version 0.7-3. 3.1

[61] B. Grün and K. Hornik, "topicmodels: An R package for fitting topic models," *Journal of Statistical Software*, vol. 40, no. 13, pp. 1–30, 2011. 3.1

[62] J. Chang, *lda: Collapsed Gibbs Sampling Methods for Topic Models*, 2015. R package version 1.4.2. 3.1

[63] T. Keys, "Using topic models package and latent dirichlet allocation to identify topics in texts." https://rstudio-pubs-static.s3.amazonaws.com/266565_171416f6c4be464fb11f7d8200c0b8f7.html. Accessed:2018-05-03. 3.1

[64] T. B. Moyers, W. R. Miller, and S. M. Hendrickson, "How does motivational interviewing work? therapist interpersonal skill predicts client involvement within motivational interviewing sessions.," *Journal of consulting and clinical psychology*, vol. 73, no. 4, p. 590, 2005. 4.2.1

Appendix A

Used Tools

A.1 Stopword List

Table A.1: Custom stopword list

a	between	former	made	own	thereupon	whither
about	beyond	formerly	many	part	these	who
above	bill	forty	may	per	they	whoever
above	both	found	me	perhaps	thickv	whole
across	bottom	four	meanwhile	please	thin	whom
after	but	from	might	put	third	whose
afterwards	by	front	mill	rather	this	why
again	call	full	mine	are	those	will
against	can	further	more	same	though	with
all	cannot	get	moreover	see	three	within
almost	cant	give	most	seem	through	without
alone	co	go	mostly	seemed	throughout	would
along	con	had	move	seeming	thru	yet
already	could	has	much	seems	thus	you
also	couldnt	hasnt	must	serious	to	your
although	cry	have	my	several	together	yours
always	de	he	myself	she	too	yourself
am	describe	hence	name	should	top	yourselves
among	detail	her	namely	show	toward	the

amongst	do	here	neither	side	towards	xxx
amoungst	done	hereafter	never	since	twelve	yeah
amount	down	hereby	nevertheless	sincere	twenty	mmhmm
an	due	herein	next	six	two	indistinct
and	during	hereupon	nine	sixty	un	youre
another	each	hers	no	so	under	dont
any	eg	herself	nobody	some	until	lot
anyhow	eight	him	none	somehow	up	didnt
anyone	either	himself	noone	someone	upon	guess
anything	eleven	his	nor	something	us	ive
anyway	else	how	not	sometime	very	wasnt
anywhere	elsewhere	however	nothing	sometimes	via	gonna
are	empty	hundred	now	somewhere	was	yep
around	enough	ie	nowhere	still	we	theyr
as	etc	if	of	such	well	ill
at	even	in	off	system	were	kinda
back	ever	inc	often	take	what	thing
be	every	indeed	on	ten	whatever	stuff
became	everyone	interest	once	than	when	theyr
because	everything	into	one	that	whence	don
become	everywhere	is	only	the	whenever	doesn
becomes	except	it	onto	their	where	mmhm
becoming	few	its	or	them	whereafter	didn
been	fifteen	itself	other	themselves	whereas	wasn
before	fify	keep	others	then	whereby	alright
beforehand	fill	last	otherwise	thence	wherein	mhmm
behind	find	latter	our	there	whereupon	theyr
being	fire	latterly	ours	thereafter	wherever	doesnt
below	first	least	ourselves	thereby	whether	yup
beside	five	less	out	therefore	which	wouldnt
besides	for	ltd	over	therein	while	makes

Appendix B

Source Code

B.1 R Studio

Listing B.1: R function topic modeling

```
#Topic modeling algorithm
#based on https://rstudio-pubs-static.s3.amazonaws.com/266565_171416f6c4be464fb11f7d8200c0b8f7.html

library(tm)
library(topicmodels)
library(tidyr)
library(tidytext)
library(dplyr)
library(lda)

#clearing Function
cleanText <- function(x) {
  x  <-
    gsub(" [ ](?=[ ])|[^-_,A-Za-z0-9,.,\\s]+",
         "'",
         x,
         ignore.case = TRUE,
         perl = TRUE)
}

#set working directory to where .txt files are stored
setwd("H:/Master Thesis Stuff/R/DataPrep/allWords_Sentence")

#read files and clean text
filenames <- list.files(getwd(),pattern="*.txt")
files <- lapply(filenames,readLines)
files <- lapply(files,cleanText)

#load document files into corpus
docs <- Corpus(VectorSource(files))

#Pipeline for stable results -> toLower, removePunctuation, removeNumbers,remove Stopwords

#Transform to lower case
docs <- tm_map(docs,content_transformer(tolower))

#Remove punctuation
docs <- tm_map(docs, removePunctuation)

#Strip digits
```

```
docs <- tm_map(docs, removeNumbers)

#create stopword list
newStopwords <- c(stopwords("SMART"),myStopwords)

#Remove stopwords
docs <- tm_map(docs, removeWords,newStopwords)

#Create document-term matrix
dtmOld <- DocumentTermMatrix(docs);
rownames(dtmOld) <- filenames

#Find the sum of words in each document and replace empty entries with 0
rowTotals <- apply(dtmOld , 1, sum)

dtm  <- dtmOld[rowTotals> 0, ]

*** LDA ***

burnin <-0 # number of omitted Gibbs iterations at beginning, by default equals 0.  1000 #set burn in
iter<-2000 # number of Gibbs iterations, by default equals 2000
thin <- iter #number of omitted in-between Gibbs iterations, by default equals iter. 500
nstart <-5 #set random starts at 5
seed <- list(254672,109,122887,145629037,2) #use random integers as seed
best <-TRUE # return the highest probability as the result
k <-6 #set number of topics

#run the LDA model
ldaOut <- LDA(dtm,k, method="Gibbs", control=list(nstart=nstart, seed = seed, best=best, burnin = burnin, iter = iter, thin=thin))

#view the top  terms for each of the 5 topics, create a matrix and write to csv
terms(ldaOut,5)
ldaOut.terms <- as.matrix(terms(ldaOut,10))
topics(ldaOut) #view the topic assignment for each document

#create a matrix and write to csv
ldaOut.topics <-as.matrix(topics(ldaOut))

#Find probabilities associated with each topic assignment
topicProbabilities <- as.data.frame(ldaOut@gamma)

allInfo <- cbind(ldaOut.topics,topicProbabilities)
write.csv(allInfo,file=paste0(Sys.Date(),"TM_All_",k,".csv"))

# write top terms to CSV
write.csv(ldaOut.terms,file=paste(Sys.Date(),"TM_All_",k,"TopTerms.csv"))
```

Listing B.2: R function sentiment analysis

```

getSentiments<- function(file){

  currFile = readLines(file)
  fnWords = data.frame(currFile)
  colnames(fnWords) <- c("fn")

  fileText <- data.frame(lapply(fnWords, as.character), stringsAsFactors=FALSE)

  # tokenization
  tokens <- fileText %>% unnest_tokens(word, fn)
  tokens <- anti_join(tokens,data.frame(word = myStopwords))
  words <- count(tokens)
  uttcount <- count(fileText)

  sentiment <- tokens %>%
    inner_join(get_sentiments("nrc"))  %>%# pull out only sentiment words
    count(sentiment) %>% # count sentiments
    spread(sentiment, n, fill = 0) %>% # made data wide rather than narrow
    mutate(words = words$n) %>%
    mutate(utt = uttcount$n)

  return(sentiment)
}
```

Statutory Declaration

I hereby declare that

- the Master thesis has been written by myself without any external unauthorised help and that it has not been submitted to any institution to achieve an academic grading;
- I have not used sources or means without citing them in the text; any thoughts from others or literal quotations are clearly marked;
- the electronically submitted Master thesis is identical to the hard copy;
- one copy of the Master Thesis is deposited and made available in the CUAS library (§ 8 Austrian Copyright Law [UrhG]).

30.10.2018, Klagenfurt

(place, date)

Nicole Hollauf

(student's signature)