ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

Master Thesis Marketing

Crystal Knows "YOU" Professors

Name Student: Louis ten Hacken
Student ID number: 343295

Supervisor: Prof. Willem Verbeke
Second assessor:

Date final version: 12-12-2016

Abstract

The purpose of this study is to investigate if Crystal is a reliable system people can use as guidance for writing letters and/or interacting with people.

The first section explores how programs such as, Crystal and Linguistic Inquiry and Word Count (LIWC), are used to discuss and present in collaboration with current literature. This will highlight how Crystal can analyse six different groups with the help of four different profiles called DISC, and how 32,203 words are analysed by LIWC.

The thesis will then identify the uses of DISC, and the reliability of Crystal. It will show that by analysing written text we can predict a person's personality. Knowing details of a person's personality will help us identify how we should communicate with people.

The results of the study indicate that Crystal is a reliable system, as proven by the group of professors. The results are showing a specific way of how we should communicate with the group professors and how to communicate with the other groups as well. The managerial implications indicate that with the help of the results, we are highlighting that we can personalize marketing messages and improve communication with all sectors of a business, through knowing how to contact a specific person.

Table of contents

1. Introduction

Every day we are using language to express ourselves. We use words to describe how we feel, and this reflects our social environment. Language is used to convince people to do things, making new friends and closing deals with our colleague. However, our language tells us more about every person than we think. The language we use can be translated to personality by making use of Crystal Knows. This program searches thousands of sources online about you and others to create a profile with a DISC personality type. DISC stands for the four styles that the model uses. Dominance, Influence, Steadiness, and Conscientiousness. These four profiles are covering a certain set of behaviours that are evident in a person's personality.

By analysing different groups of people and looking into their personalities, we can see how people use personal branding to engage with others through the use of online platforms (Labrecque, 2011). As marketer you can use this information to personalize ads, emails, and identify how to negotiate with your customers.

Linguistic Inquiry and Word Count (LIWC) is a text analysis program that counts words and separates them into psychologically meaningful categories. With LIWC we have the ability of detect a variety of languages between different groups of people. Using LIWC in combination with Crystal will help us find how

reliable the outcomes of using Crystal are.

An expert in researching what languages is telling about a person is James Pennebaker. His book "*The secret life of pronouns what our words say about us*" (2013) explores how the most commonly, smallest, most forgettable words can serve as windows into a person's behaviour, thoughts and emotions. The book explains how the words, pronouns and articles a person uses are linked to their personality, honesty, social skills, and intentions. "*how everyday language reflects basic social and personality processes*" (ALLC, 2013).

Research into writing language, stylometry, dates back to 1851. At this time it has been mostly used to prove who the author was (Homes, 1998). Lately the focus of language is more on the emotional and health of a people than the cognitive and emotional processes that are involved in it. Multiple research papers indicate that talking and writing about your problems can improve a person's health. Which shows again that analysing our language, by making use of a computer program shows us a "fingerprint" of the words people use (Pennebaker & Graybeal, 2001) (Davison, Pennebaker, & Dickerson, 2000).

Contacting people by using their personality traits is the next step in marketing and in the personalisation of e-commerce. Large data sets of text and online profiles, makes it possible for marketers to analyse their customers. This sort of data mining, such as the research I conducted, are becoming more

popular over the last couple of years. Data mining gives us the opportunity to learn from our customers. It also gives us the ability to turn customer data into customer knowledge (Linoff, 2011).

For this research we analysed, every word of every person in the data set. By analysing every word, we created a large data set that we can apply the outcomes not only for the people that we analysed, but we can make generalisations from the information. Also by analysing every word, people can make assumptions on completely different topics without changing the outcomes.

The following chapters will first explain the research and how Crystal Knows and LIWC help us reach these outcomes (e.g., LIWC2015). These results will show us how to communicate with a certain group of people. It will also highlight the differences between groups and how to use these differences in marketing. In the last chapter we aim, to determine whether Crystal is a reliable system that people can use as guidance for writing letters and/or interacting with people.

2. Thesis Statement

Through analysing personality traits of different groups, we can establish how to interact with certain groups of people and in particular, with professors. We will also explore how people use personal branding in their language. Personal branding can be seen in online profiles, in a way that people manage online profiles so that other people will have the "right" judgment of the content they post. Finding out how written text plays a role in personal branding is a problem that fits research in marketing.

Looking at the language your customers use will allow you to interact and communicate with them more effectively (Ng, 1993). Through examining how customers communicate, we can establish target customers based on the language use. By knowing how to communicate to your customers we can establish how to steer marketing efforts towards them through speaking and writing (Ramsey, 1997).

3. Approach and Methods

Based on suggestions and discussion with my thesis advisor, we found a interesting program called Crystal Knows. This research will test the data generated by Crystal Knows. Through choosing categories that substantially differ, we can test the content. Hence consider this project as an exploration rather than a replication.

The data collection is based on six pre-determent groups. The six groups are CEO/CFO, team sport, individual sport, politics, medics, and professors. For every group we analysed every word of every person. The reason for these six specific groups is to predict the different motivations between these groups. Intuitively, sports people love teams, medics love status, individuals people love to excel and test their own limits. I assume that these groups contrast substantially in language used to express their motivations differ so much. In the end my goal is to establish, whether Crystal is a reliable system that people can use as a guidance for writing letters and/or interacting with different kinds of people.

To clarify this, we have figure 1 which shows us a schematic overview of what we expect. Within a group, every person has their own personality, but because they share a common motivation, I expect that their language expressions   have   similarities.
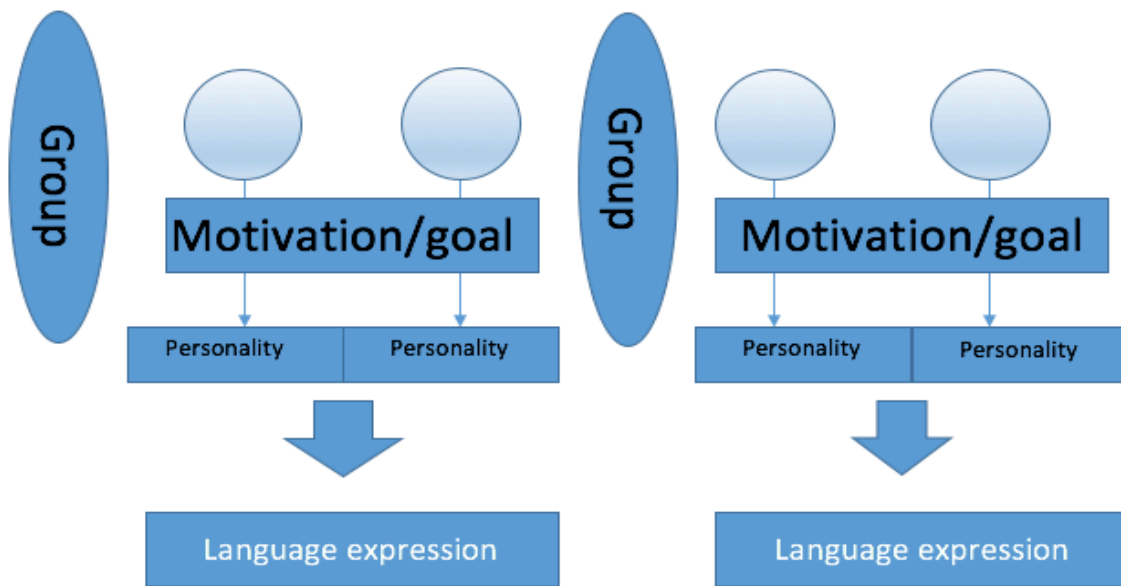
Figure 1. The psychological drive behind the personality traits.

Within each group the collection of people is based on a random search. The only restriction is that the person needs to have a LinkedIn account. The importance of having a LinkedIn account will be explained later when we describe Crystal Knows. Every group consist of 50 people, working for different companies. For the sports category, they not only need to have a LinkedIn account they also need to perform on the highest level/earning their money with sport, only then they are categorised as sport. The data set it self is larger, however as not every person has a LinkedIn account we could not use all the profiles. There are also some exceptions as a person may have a LinkedIn account but, Crystal knows could not find them.

3.1 Analysing step by step

Throughout this section I will explain the step by step process of the

research. Initially, I started by establishing my six different groups. Within these groups I randomly choose 50 or more people. I then started to analyse every person by using Crystal. For every group, all the outcomes of Crystal were put into one large data set. These outcomes were then put into LIWC which analysed every word of every person. This is around a 100 to 110 words for every person that were analysed. This gives us a total set of 31.500 analysed words by LIWC. The outcomes, of LIWC which I clustered by group, are then analysed by SPSS. It is important to remember that we got all the data from Crystal and not from the individuals themselves.

3.2 Crystal knows

"*Crystal analyses public data to detect your DISC personality type without a test*" (Crystalknows, 2016). Crystal starts by examining all the things you have written on the web and that can be viewed publicly. It looks primary into social media profiles. This is also the reason why people also need a LinkedIn account. LinkedIn is written by the person them self, which makes it an ideal indicator for Crystal. In conjunction with LinkedIn, it utilizes other social media profiles and looks at your writing style and sentence structure. Once it has explored what the individual has published online, it processes what others have written about you on these social sources.

Using these data points as a start, Crystal categorizes an individual as

possessing a certain personality traits making use of the DISC personality tool. This personality tool is basically dividing the people in four categories. When Crystal find a person it produces a profile with a short summary, as shown in figure 2.
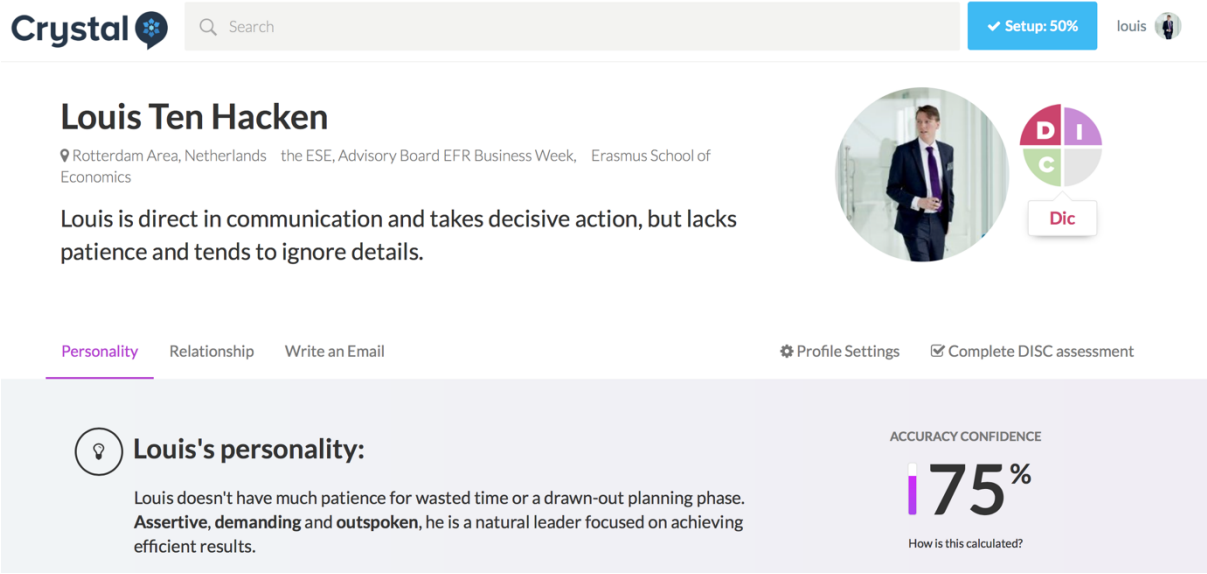
Figure 2. Crystal summary, telling us the DISC outcome and your personality.

Part of this profile is the personality by DISC type and the certainty about the profile. The percentage of certainty of a profile depends on the available information online about a person. These DISC types explain a lot about how to communicate with one and an other.

DISC is a personality assessment tool created by Marston. He explores this in his book *'Emotions of Normal People' (1920)*. To be clear, Marston did not develop an assessment or test. However, he came up with the model of DISC, which others later used to develop the assessment. Crystal is using DISC as it is

more specific as it can have a mix of different types and subtypes.

Figure 3. DISC overview

A lot of companies are using DISC in many different ways. Most common way it is used is through people filling in a questionnaire, based on the answer it separates individuals into one of the four categories. Two researches show that the primary personality and behaviour patterns of entrepreneurs can be found with the DISC assessment tool. First research, is about small business entrepreneurs. It shows that two categories had a higher propensity to start a business. This tells us that being an entrepreneur can be seen in the DISC assessment test (Krueger, 1998). Second research, find that most female entrepreneurs are from the dominant personality category. Which indicates, that for females having this personality, it makes it more likely to become entrepreneur (Krueger, 2000).

In table 1, we see an overview of the four types that DISC is using. For this research people don't fill out a questionnaire to determine what type of personality they have. It is Crystal who determines the type of personality a

person has based on what you have written online and what they can find about you.

Type	Short description
Dominant	Confident, aggressive, result-oriented, and direct. There communication to others will be very short and to the point.
Influential	Open, outgoing, persuasive, enthusiastic and social personality. In communication with others they are very talkative and think out loud.
Steady	Sincere, dependable, and steady. In Communication an S-personality will be calm and reserved.
Calculating	A C-personality values rules, quality, objectivity, competency, and autonomy. In communication to others they will rely on logic and accuracy.

Table 1. Overview of the four different types in the DISC model.

To answer the main question of whether Crystal is a reliable system that people can use as a guidance for writing letters and/or interacting with certain types of people. We divided the question in sub hypothesis. The first helps us to check on the diversity of the groups.

Hypothesis 1:

Disc looks at personality traits. Given the diversity of the population: team athletes, individual athletes, medics, politics, professors or directors in firms we expect that the populations will differ substantially.

Crystal is not only giving us a DISC personality it shows us what comes naturally, by what a person is motivated. How we should communicate with this certain person by e-mail but also when speaking, and how we should work with

this person.

For this research we take 4 of these groups; speaking, emailing, working with, and comes naturally to. These four groups all consisting out of 50 people or more and are put in one document, so that we are able to analyse them with Linguistic Inquiry and Word Count (LIWC). See exhibit 1 for an example of what we analysed with LIWC. Good to remember is that Crystal is using a disclaimer on their website to tell us that the results they provide are a "best guess" about a person.

3.2.1 What others think of Crystal?

When looking on the web, I find some interesting views on Crystal. In general people find it pretty accurate what Crystal is finding about them. At the same time people are asking if this is not going to far. A program that can tell us how to communicate with others is that not getting creepy?

Question likes this are coming back in all the reviews that I can find over Crystal. "*Crystal knows best … or to much? The disconcerting new email advice service.*" (2015), "*Does Crystal know you better than you know yourself?*" (2015), "*Personal branding to the max want Crystalknows 'kent' je.*" (2016), "*Stalk everyone you know with this eerily accurate app that tells you how to talk to people.*" (2015) and, "*Here's What Happened When I Let an App Dictate How I Wrote Emails*" (2016). All these blogs/news pages are talking about two sides of Crystal, one it is brilliant that it can be so accurate when looking for how to

communicate with a person you don't know that well. But secondly they always come back to the privacy part. As is this not going to far in predicting people by analysing their social media profiles? In figure 4 you can see some quotes of what people say about Crystal on Twitter.
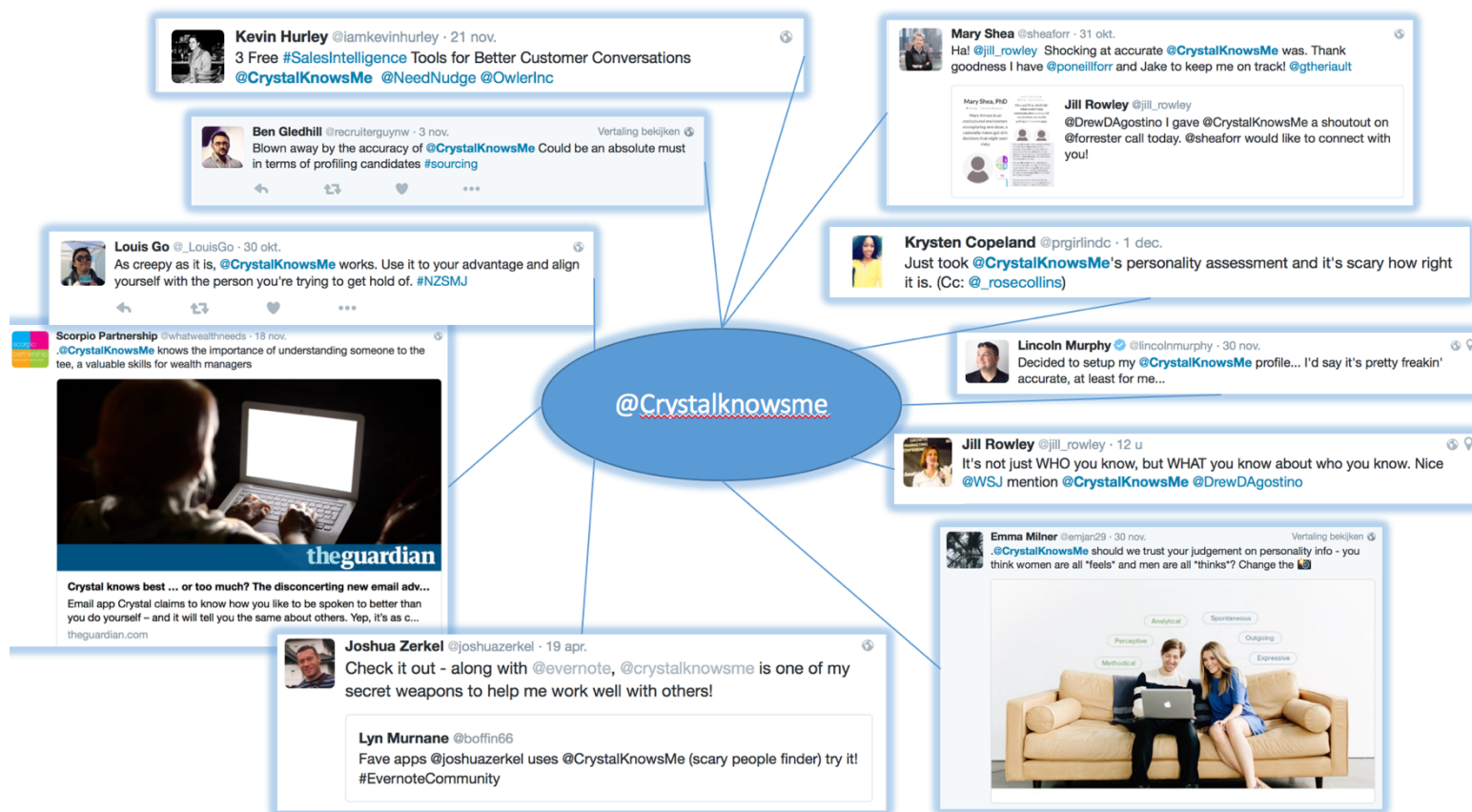

Figure 4. Quotes from @Crystalknowsme

3.3 Linguistic Inquiry and Word Count

Linguistic Inquiry and Word Count is a text analysis program that counts words and put them in psychologically categories. It reads a given text and counts the words that reflect different emotions, thinking styles, and social concerns (Pennebaker, 2015). The LIWC program has two features, the processing

component and the dictionaries. The processing part goes word by word well the dictionaries are comparing it with the dictionary file. Based on this comparison they put the words in a category. LIWC used a specific method to updated and check all the categories[1]. In 1997, 2007 and again in 2015 LIWC had a revision well the dictionaries where updated and streamlined. For this research we made use of the LIWC2015 version with the standard English dictionary.

LIWC shows us the percentage of words that are categorized in a certain group. All the personalities that we got from Crystal are analysed by LIWC. Which gives us an overview for every group and for every person. The more words they use in a specific category the higher the percentage score they receive for this category. The total amount of words analysed over all the groups by LIWC is 32,203. This brings us the the second hypothesis:

Hypothesis 2:

Crystal only analyses what professionals write and then make statements how to speak/write to them. Hence this is a mirroring effect. Therefore, I expect that Crystal (as analysed by LIWC) will also show differences in the words people across different professions should use to address them.

---

[1] For more information about the judge of the categories within LIWC; *"The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods"(Tausczik, YR., & Pennebaker, J.W. (2010)).*

3.4 Literature review

In the next part we will shortly go into the work of James W. Pennebaker and the relevance of the outcomes for this research.

In the paper "*Language Style Matching Predicts Relationship Initiation and Stability*" (2011) Pennebaker and his team explore the most overlooked part, that people talk with each other. Similarity in personality traits, values and interests is known to predict mate selection (Watson & Klohnen, 2004). Conversations often serve as the basis of attraction. By making use of language style matching (LSM), they predict outcomes for relationships. The conclusion of the paper is that LSM predicts relationship stability (Ireland & Slatcher, 2011). Using this means that if we match our language style to someone, that he or she is more likely to like you. Translate this to a brand or advertisement means that if we can match the language of our target audience that it is likely that they like your brand more.

The paper "*Patterns of Natural Language Use: Disclosure, Personality, and Social Integration*" (2001) explains multiple researches about how writing and speaking is having influence on our physiological changes and health. It started 15 years ago when the first research started by finding out that, when people write about their traumatic or emotionally experiences that it could affect mental and physical health. It is found that it is not about blowing off steam by writing it down. But it affects the way people think about themselves. This is

where LIWC comes in, as it can be use to explore how language is used during the writing. The first research they conduct was if emotion words where related to health. As LIWC can count the amount of emotional words. What they found was that cognitive words where a predictor of health. Another important finding in this paper is the implication that linguistic styles can be considered as markers of personality. And that people have their own styles of language use and that the use of everyday language correlates with health and social behaviours.

James Pennebaker has wrote the book "*The secret life of pronouns. What words say about us*" (2013). In this book he explains a couple concepts as well as how Tweets and Facebook posts can tell us about the personality of their authors. Overall he shows us that function words can reveal personality traits (ALLC, 2013).

The article "*The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods*" (2010). Reviews different methods of text analysis. And describes how LIWC is created and validated. It also shows us a problem of testing validation of LIWC. As people are never telling exactly the same thing twice. This makes test-retest reliability not possible. The article talks about that researchers are trying to figure out the "real" person. Making use of priming strategies and various imaging techniques such as functional MRI. They also consider analysis of language to do the same. Especially function or style words.

But they also point out that computerized language measures are still crude. They ignore context, irony, idioms, and sarcasm (Tausczik & Pennebaker, 2010).

The paper "*Psychological aspects of natural language use: Our words, our selves*" (2003). Gives a brief review of several text analysis programs, summarized evidence that links natural word us to personality, social and situational fluctuations, and psychological interventions. This article explores the methods and recent findings on word use rather than language per se; the style in which people use the words rather than the content of what they say. There are three broad topics that this paper is focussed on. 1) The way researchers have studied the ways people naturally use words. (natural, relatively open-ended responses to questions, natural interactions, and written or spoken text). 2) Exploring recent findings linking word use to individual differences. 3) Links between word usage and social or situational differences and how we can use words to psychological change.

Main points: capture a maximum of words in a given text (Biber's approach, LIWC) or concentrate on only some linguistic aspects (TAS/C, need-achievement). Both get quite different picture but they are equally valid.

Markers of linguistic style are generally associated with relatively common words such as pronouns and articles. Many of the more content-heavy words-nouns, regular verbs, and modifiers – have not yielded many consistent social or

psychological effects. This may reflect the fact that linguistic content is heavily dependent on the situation or topic the person is instructed to think or talk about (Pennebaker, Mehl, & Niederhoff, 2003).

Can language use reflect personality style, that is the question that the paper *"Linguistic Styles: Language use as an individual difference"* (Pennebaker J. W., 1991) tries to answer. Pennebaker and his team look at multiple researches, and languages. The data from their research demonstrate that the way people express themselves in words are reliable across time and situations. The conclusion; modest effect sizes, the data suggest that linguistic style is an independent and meaningful way of exploring personality. (Pennebaker & King, 1999)

These six papers are all covering more or less the same. Language as indicator for how we are as a person! The goal of this paper is to find out if Crystal is a reliable system to use as guidance for writing to people. The articles above give more insights about how language is used in different research and how LIWC can help us to find language expressions within the different groups. Assuming that people with the same motivation use more or less the same language expressions makes it possible to use LIWC on the outcomes of Crystal. Crystal gives us personalities based on a search of public data, knowing that linguistic styles can be considered as markers of personality and as an

independent way of exploring personality. This gives us the indication that by looking into the linguistic part of LIWC should help us find a relationship between personality and different linguistic categories. Based on the suggestion that I think that professors are a unique group within the six groups, I came up with the last hypothesis that will help us answer the main question.

Hypothesis 3:

Professors are a unique group in DISC: they are very strategic which indicates that they have much freedom to do what they want in life compared to other professions in the study. Therefore, we expect that Crystal will use special words (as analysed by LIWC) that differ substantially from the other groups.

4. Results

The results are divided in two parts; first we take a look at the analysed data by LIWC. After that we take a look at the outcomes of Crystal and what these outcomes say about the different groups. The outcomes of Crystal are analysed with LIWC2015. First we start with looking only at a specific amount of categories. We are not interested in all of them, which leaves room for more research.

When looking at the different graphs, the X-as shows the six different groups while, the Y-as shows the percentage of words you should use when communicate to a group. In this research we put more focus on the group professors as they show more deviating outcomes.

4.1 LIWC results

LIWC has four groups, we only focus on psychological processes in this category. First looking at positive emotion, negative emotions and sadness. In exhibit 2 you can find the outcomes of SPSS.
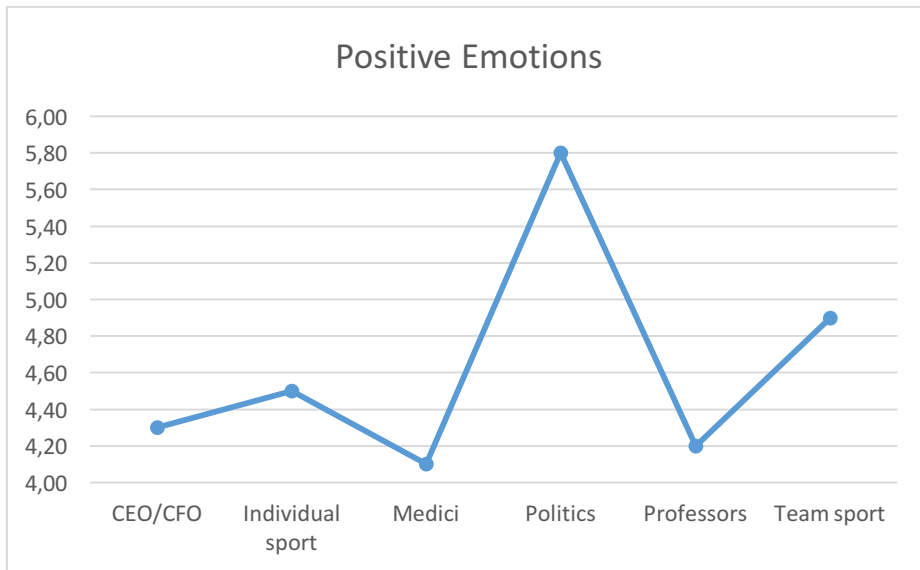
Figure 5. Scores of positive emotions

Looking at the first graph (figure 5), shows a lower score for positive emotions but a non-significant score. Within figure 5, the group politicians attract attention with its high peak.

The second graph shows us a lower score of negative emotions for the group professors, with $a = 0,01$ (figure 6). When examining these graphs, you need to remember that it informs us, how we should communicate with a specific group and not how they should communicate with us.

This graph provides insight into negative emotions and a low score indicates that we should not use negative emotions in our communication with professors. As we can see in the graph, professors received a low score.

Figure 6. Scores of negative emotions
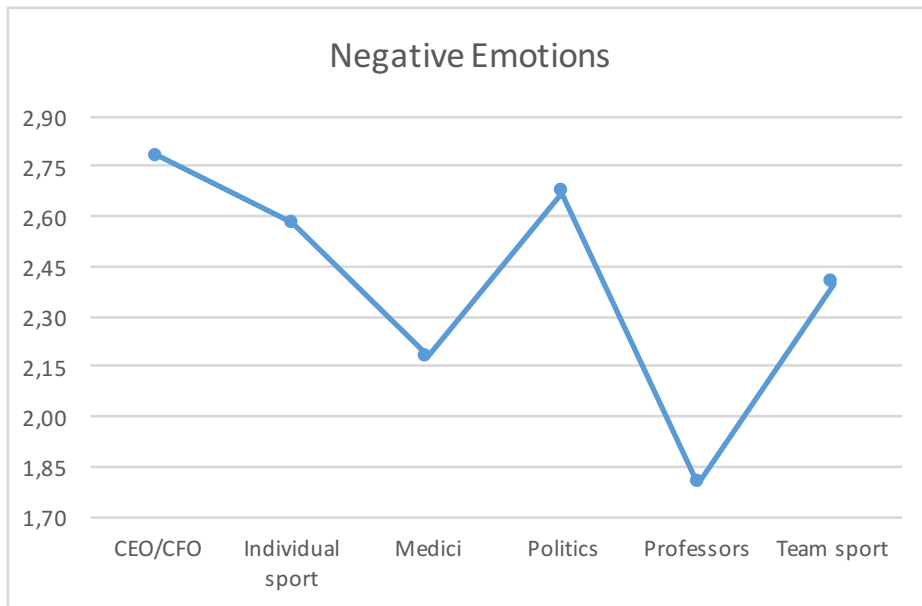
The third graph represents sadness $a = 0,05$ (figure 7). Just like negative emotions, the group of professors received a lower score, which tells us that we should not use incorporate sadness when communicating to a professor.
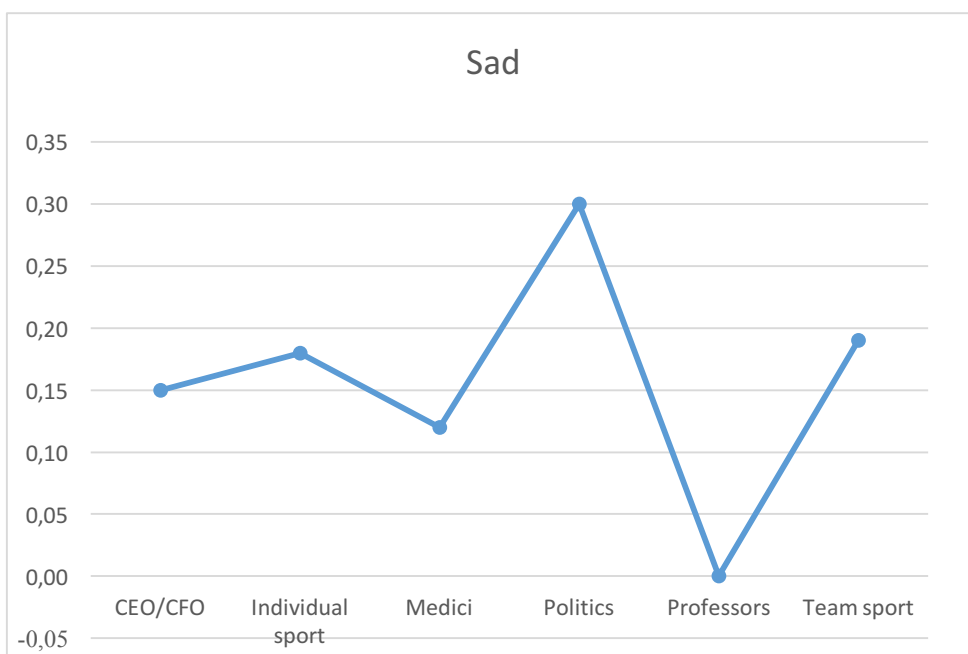
Figure 7. Scores of sad words

In these three graphs we see that when, communicating with professors we

should not use "sad" words and we should not use too much negative emotion in our words. This indicates that we should use positive communicate with professors. For the group of politics, you can say when communication with politicians use a lot of emotion in your communication due to the higher scores shown in all the three graphs.

Now reviewing the cognitive processes, we examine at insight($a = 0,00$), causation($a = 0,00$), certainty($a = 0,00$), and differentiation($a = 0,00$). All of the groups show significant, in exhibit 3 these outcomes are demonstrated. But what does it mean for our communication with professors?
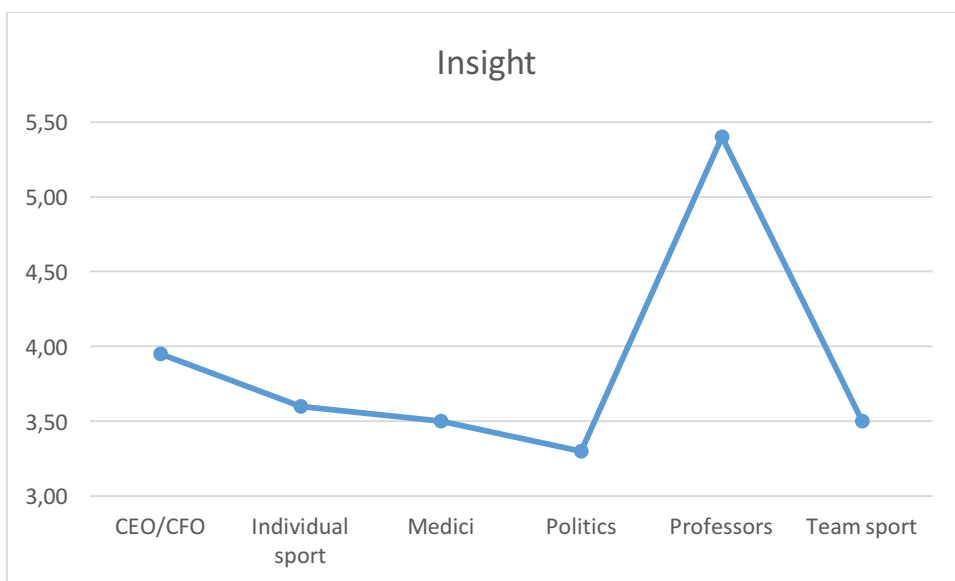
Figure 8. Scores of insight.

In the graph showing insight (figure 8), we see that professors score higher than any other group. Which means that when we communicate with professors it is important to always think before saying something, and use data to support your point.
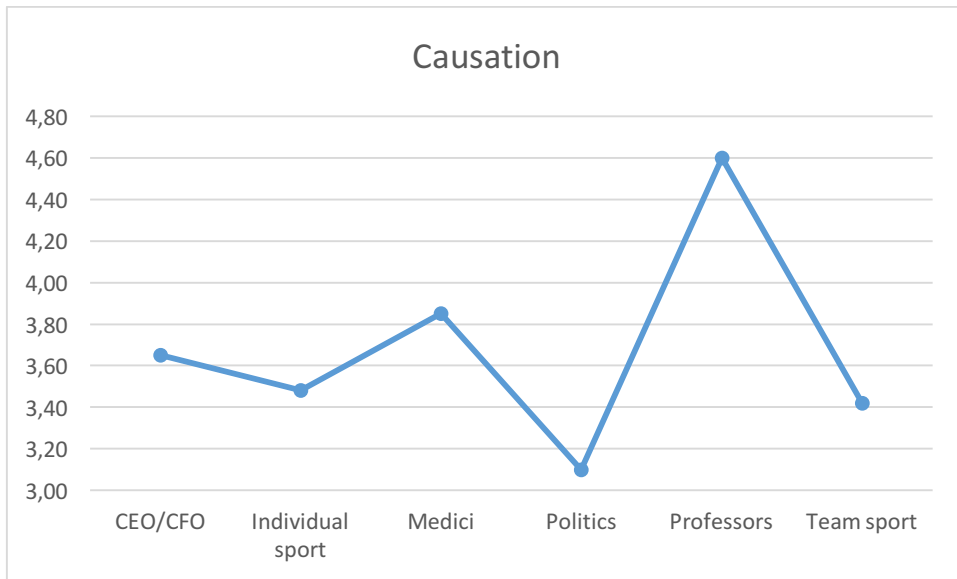
Figure 9. Scores of causation.

Professors also score high on causation telling us that when communicate

to professors you should look for causation.
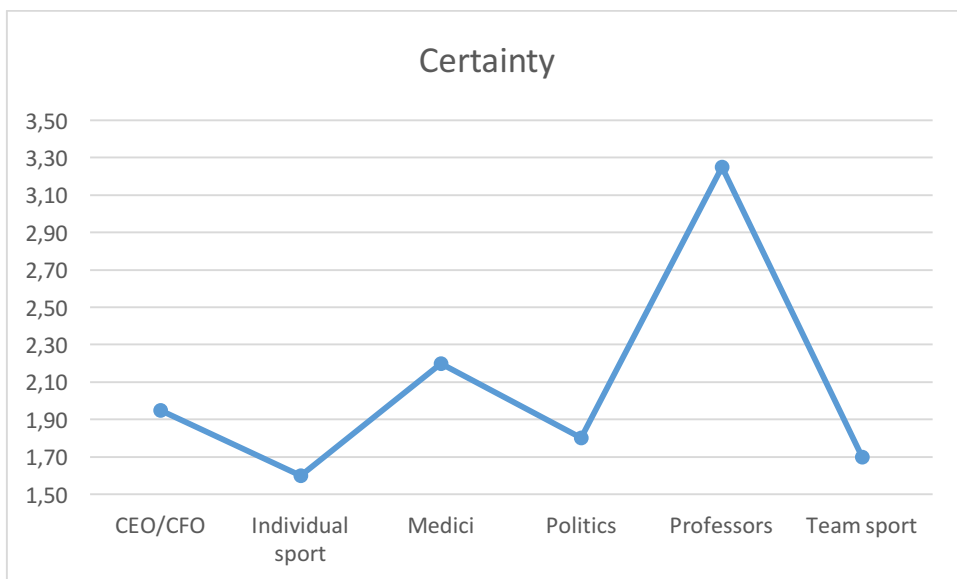
Figure 10. Scores for certainty

Professors score higher on certainty and lower on differentiation, keeping

this in mind when communicating with professors a situation is either true or

false, but can not be both.
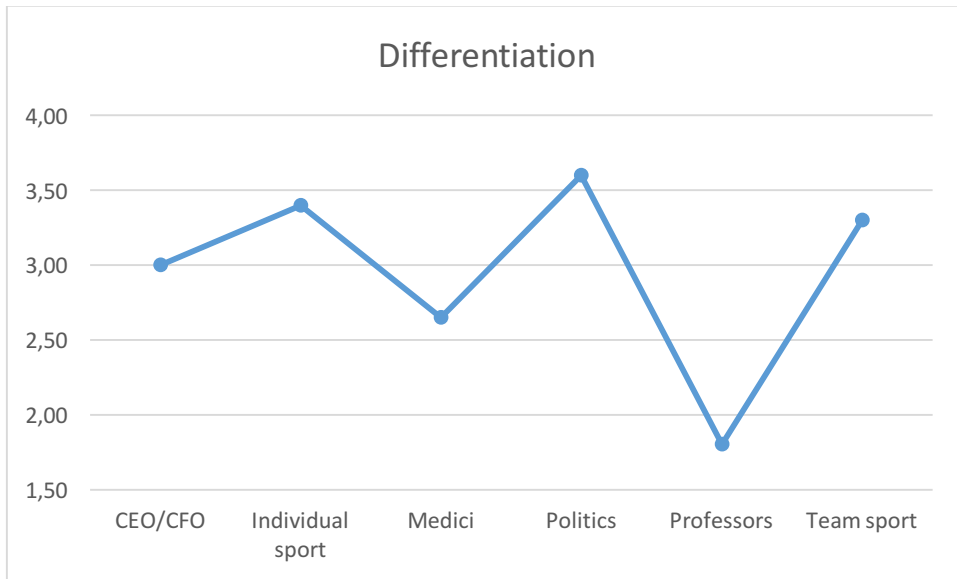
Figure 11. Scores for differentiation

In four categories; insight, causation, certainty, and differentiation we can see that the group of professors are a good example of a paradigm. Interestingly, within these four categories the professors always score completely differently compared to the other groups. There is no other group that communicates in the same way.

When examining what drives a certain group, we analyse the categories affiliation ($a = 0,00$), reward ($a = 0,00$), and risk ($a = 0,069$).
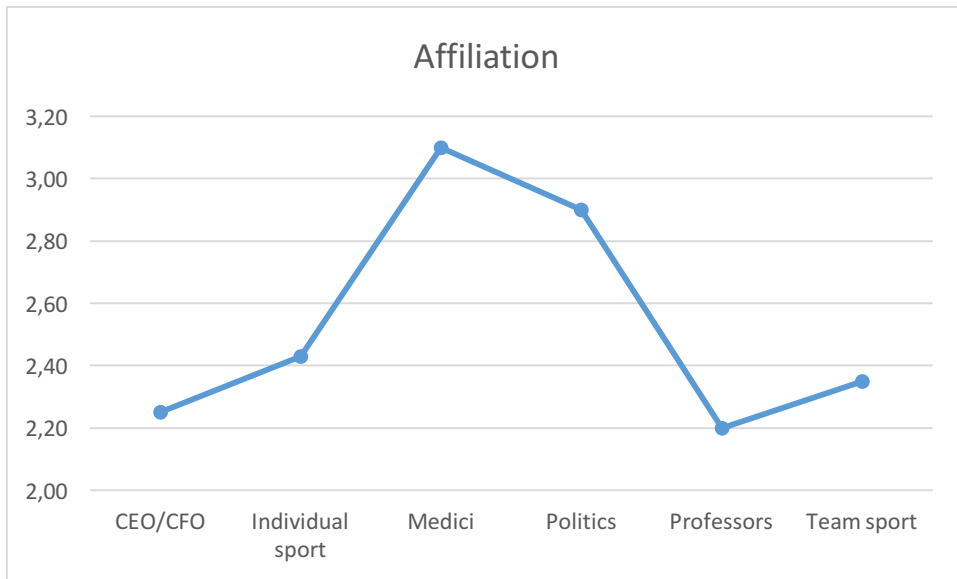
Figure 12. Scores for affiliation

For the category affiliation, we see a lower score for professors together with the groups of CEO/CFO and team sports. This indicates that when communicating with any of these groups, you should not presume or mention that you are friends with him or her.
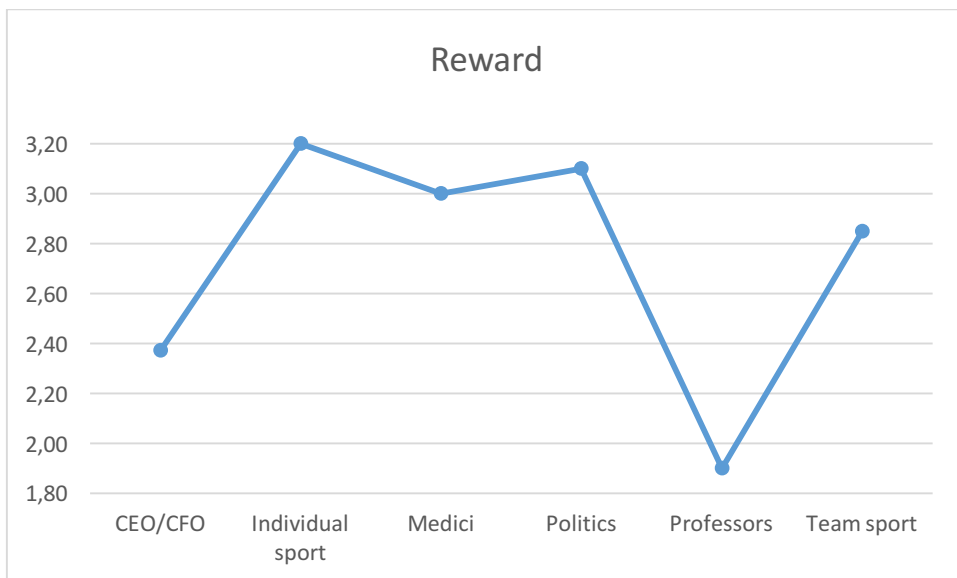
Figure 13. Scores for reward

The category reward highlights that when communicating to professors you should stay objective; they are not looking to be rewarded. Its the nature of

their work in science and the research, where they don't work to be rewarded. At the same time professors are driven by risk, so when communicating with them you should mention risk related words, like danger, doubt, risky, etc.
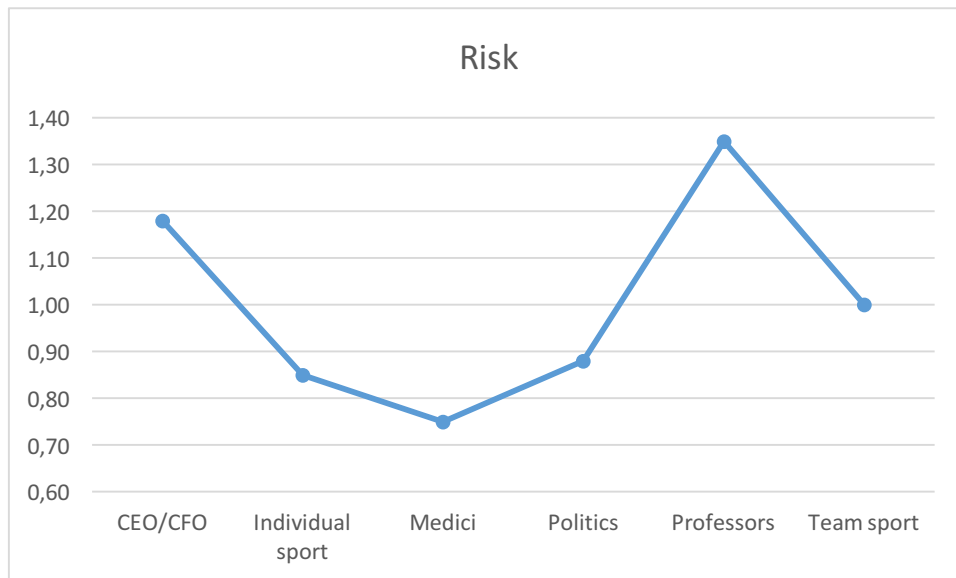
Figure 14. Scores for risk

Through examining these three categories, we can conclude that when communicating with a professor it is important not to mention that you are his or her friend, stay objective and do not reward them as, they like to see risk is being taken! With the category affiliation, it is shown that they prefer a strict and clear line between friends and work. This indicates that it is important to not get too personal and enter their comfort zone.

Last category we are looking at, are the personal concerns of the six groups. Looking at work $(a = 0,00)$ and leisure $(a = 0,00)$ both with a significant score see also exhibit 5.
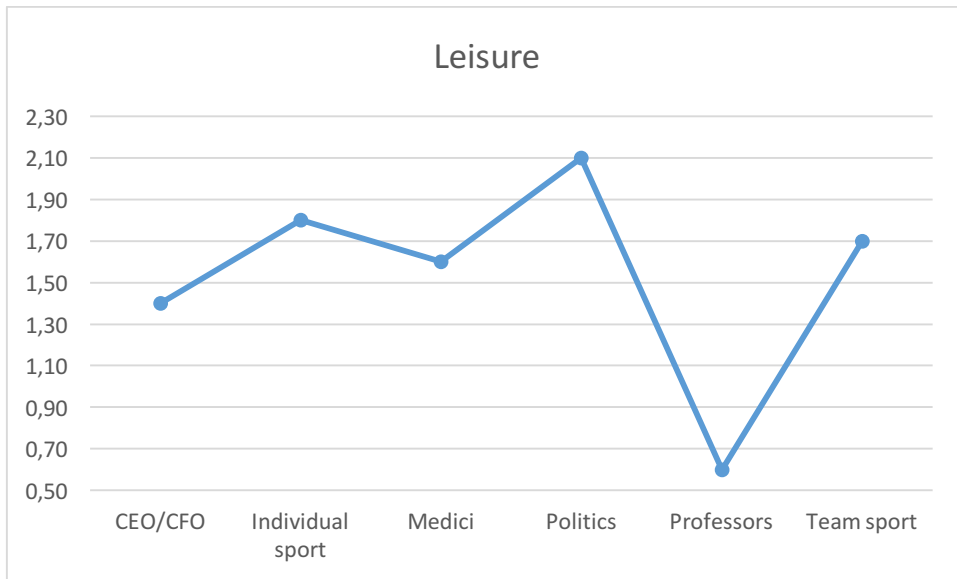
Figure 15. Scores for leisure

For these categories, we can gain a clear understanding into how professors would like to be communicated with. A low score for leisure (figure 15) and a high score for work (figure 16) indicates that you should not mention or talk about leisure, but instead focus on work.
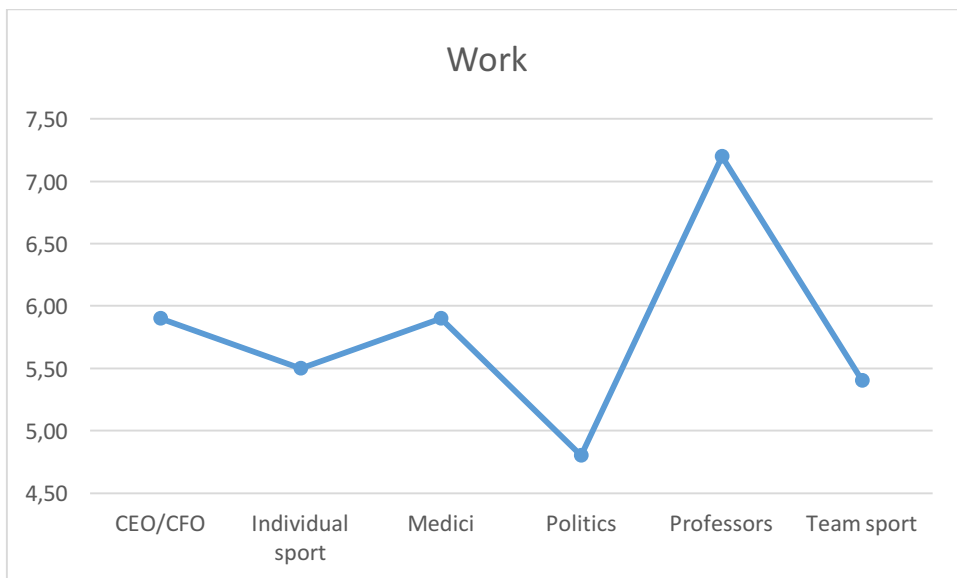
Figure 16. Scores for work

From here, we look at the second part of the results. These results show us the outcomes from Crystal Knows. Crystal makes an individual DISC profile for every person. In table 2 we have counted the DISC profiles. When the DISC profile existed out of a combination, then we take the primary type.

	CEO-CFO	Team sport	Individual sport	Politics	Medici	Professors
D - Dominant	16	11	10	6	2	6
I - Influencing	7	8	12	1	2	0
S - Steady	14	12	13	19	13	11
C - Calculated	6	3	2	1	9	25
0/?	7	14	13	23	24	18
total	50	48	50	50	50	60

Table 2. Overview DISC profiles from Crystalknows.com

In most of the groups, we see one or two primary types are the most common. For the group politics, 19 out of 50 have steady as their primary type, same for medics with 13 out of 50. For the group professors, we see 25 out of 60 having calculated as primary type, followed by steady 11 out of 60. For the group individual sport, we see that there are only two athletes with calculated as primary type, while the other types are almost equally distributed over the other athletes. 10 out of 50 dominant, 12 out of 50 influencing, and 13 out of 50 steady. Same happened in the group team sport. Only 3 athletes with calculated

as primary type. While dominant scores 11 out of 48 next to 12 out of 48 steady. In the last group CEO-CFO we see 2 primary types dominant 16 out of 50 and steady 14 out of 50.

Within every group, there is the problem that Crystal could not find a DISC profile. This shows up as a zero or question mark on the sheet. Especially for the group medics and politics, where almost 50% of the group Crystal could not be found as a primary DISC type. This should be considered when looking to the other primary DISC types.

These primary DISC types provide an insight into the different groups and answers one of the questions we want to focus on. It highlights the way we should communicate with a specific group of people, for example the group professors. 25 professors have calculated as primary type. This tells us that they are "*Very analytical, gravitating towards process, structure, and rules. Professors are intensely sceptical and use logic to objectively make decisions, rather than being swayed by emotions*" (2016). We can ask ourselves does this makes sense for the group? Yes, as a professor is driven by logic to explain theories, as this is a crucial part of the job. In addition, it also gives us valuable information in how to contact and communicate with a professor.

Let's have a look at the group CEO/CFO, dominant and steady are the most primary types. What does this tell us? "*CEO's are direct assertive, independent,*

and decisive. The tend to be intense competitor that thrive with ambitious goals and challenges, preferring action over analysis when they need to complete a task" (2016). And for those CEO's with steady as their primary type, "*Reserved, extremely supportive, consistent, and loyal individuals. These CEO's are known for their empathy and compassion for others, which helps them foster close long-term relationships*" (2016). It is evident that there are two different kind of leaders. The first type of leader is direct, with ambitious goals and is up for challenges. While the second type of leader is the more reliable boss, who adopts a more supportive roll with their employees. In my opinion, this makes sense for CEO's, as there is not one way in how you should lead a group of people.

In the group medics, 13 out of 50 have steady as their primary type followed by 9 out of 50 with calculated. Remember, that this group has a large number without a DISC profile at all. If we only look to those with a DISC profile, the results would be 50% steady as primary type. Looking at a steady personality type tells us, "*That medics are reserved, innately supportive, consistent, and loyal. But also calm patient and thoughtful in all that they do*". For those medics with a calculated profile "*analytical, gravitates towards process, structure and rules. Being intensely sceptical and use logic to objectively make decisions*" (2016). Ideal characteristics to see in a doctor as they are helping people in need.

For both types, we see a personality that thinks about their actions before doing something. Also working in a structured environment will help doctors perform better in complicated situation that can be the difference between life and death.

Politics are having a big group of people with steady as primary type. For this group also counts that Crystal could not find a DISC profile for 46% of them. Steady as a personality type gives us, *"That they are reserved, innately supportive, consistent, and loyal. But also calm patient and thoughtful in all that they do"* (2016)*.* In my own opinion not directly the personality I see with a political person. But at the same time they are chosen by people to represent them, which can explain the loyal, and supportive factor within steady.

Within the group individual sport, we can detect 3 major personality's types. Dominant, influential, and steady this can be explained as athletes are coming from a lot of different backgrounds, from ice skating and swimming to playing golf. For those with influential as primary type, *"These athletes are very confident and fun to be around, individuals with love for social settings, excited to explore fresh ideas and begin new projects"*. Dominant as primary type *"Direct, assertive, independent and decisive. These athletes are intense competitors that thrive with ambitious goals and challenges. They prefer action over analysis when they need to complete a task. They push harder then other*

more passive personality styles". Last primary type we see at these athletes is steady *"These athletes are more reserved, supportive, consistent, and loyal individuals. They rely more on assertive acquaintances to take control in a situation"* (2016). These 3 primary types most likely are closely related to the kind of sport someone is doing.

Within the group team sport we see a similar picture as with individual athletes. Three personality profiles are almost equal distributed. For team sport we see dominant and steady a bit more then influencing. The athletes in this research are mostly soccer players or field hockey. So we can not say that the difference in personality comes from a lot of different sports. What I expect is that teams and athletes perform better when they are with different personalities around them. That can be in their own team of within their team of support staff. Other reason for the three personality groups is that personality probably is not predicting if you can be a good athlete. In the end athletes need an optimal construction of the body and their personality is probably less important.

5. Hypothesis

Statistical tool IBM SPSS will be used for analysis and all of the outcomes will be interpreted and discussed based on these results.

Looking at the outcomes, they do provide significant evidence to support

the first hypothesis: "*Disc looks at personality traits. Given the diversity of the population: team athletes, individual athletes, medics, politics, professors or directors in firms we expect that the populations will differ substantially*". Looking at the results we see that the groups do differ substantially on most of the categories. See also exhibit 2 till 5. Therefore, we can support hypothesis 1.

Looking at hypothesis 2 *"Crystal only analyses what professionals write and then make statements how to speak/write to them. Hence this is a mirroring effect. Therefore, we expect that Crystal (as analysed by LIWC) will also show differences in the words people across different professions should use to address them".* First we have seen that crystal is making a mirroring effect by looking how people communicated and based on that it tells us how to communicated back. What we see in the results is that depending on your personality we should speak to you in a certain way. What we also see is that not every profession is made out of 1 main DISC type. But overall we can see in the results of LIWC that people use different words across different professions. And thereby, we support hypothesis 2.

The last hypothesis is about one specific group the professors. *"Professors are a unique group in DISC: they are very strategic which indicates that they have much freedom to do what they want in life compared to other professions in the study. Therefore, we expect that Crystal will use special words (as analysed by*

LIWC) that differ substantially from the other groups".

That the group professors are a unique group can be seen back in table 2. There we see that professors are the only group with a large proportion of personalities with calculated. With LIWC we see that professors differ substantially in a lot of categories. As example you can look at the category personal concerns, and measure the amount of words used in the category work and leisure. What we see in figure 15 and 16, is that the group professor differs substantially from the other groups. See also appendix 2 till 5 for all the significant scores. Seeing all these difference for professors compared to the other groups through the outcomes of LIWC tells us that we can support hypothesis 3. Crystal is using special words that differ substantially from the other groups.

6. Conclusion & Discussion

The previous chapter highlighted the results of the study and in the current chapter those results will be discussed, including their implications for marketers and companies.

The results showing us a difference between the groups, with the most interesting group professors. Not only finds Crystal that they are different compared to the other groups. Also LIWC shows us that the personality of the group professors is different then the other groups.

We see that the personality calculated from DISC describes the professors as analytical, sceptical and that they use logic to make decisions. This is the same outcome in LIWC where the professors score high on insight and causation see exhibit 3. Its important to remember that the outcomes are telling us how we should communicate with a certain group. It does not tell us how a specific group is communicating back to us. We can make some assumptions but the goal of this research is to show how you should communicate with a group and in specific with professors.

Communication can help us within the marketing field, you can influence people by communication in a way that fits their personality. Therefore, you could say, analysing language is the next step in personalisation of marketing messages.

Combining the outcomes of Crystal and LIWC shows us some interesting conclusions. With a focus for group professors, we can say the following about them. When communicating with professors, stay focussed on work related topics and not on leisure. Get right to the bottom line and use data to prove a point in your story. Stay objective, they are not looking to be rewarded. They like risk taking, so exploring of new areas/ideas which fit with having a job in research. All the communication should be in a rather positive way; they don't like the use of negative emotions in their conversations. When we know all these

communication preferences of a group, it can be used in influencing an individual within this group. Personalize the communication to your target group will give you an advance compared to regular communication. For marketing purpose, you should think of messages that are written in such a way that it is matching your target group.

To answer the main research question, it can be concluded that Crystal is a reliable system for guidance in writing or interacting with people. The study indicates that based on language we can tell how we should communicate to a person. By controlling the results with LIWC we can say that Crystal is a reliable system for guidance in writing and/or interacting with people.

This conclusion can be used for marketing purpose in a way that we can personalize ads not only on what people are looking for but also, written in a way that is matching to their personality. Crystal can have an interesting impact on how we use communication now a day. By knowing which communication works with what kind of people, makes it possible to influence people with communication without that they know that we are influencing them. We have seen that the way a person communicate is like a fingerprint, it is unique for every person. It is an interesting new part of looking to marketing and can be used in all kinds of marketing, from more personalized advertisement to websites that use a specific layout and communication based on what kind of

personality you have. In the end this kind of "personalized" communication will be used to influence the customers.

7. Limitations

In this last chapter, a few limitations are mentioned. From this limitations, a recommendation for future research is given.

Only focussing on the outcomes of the group professors with LIWC brings it limitations. As this leaves room for further research into what languages tells us about a person. And specifically about the groups that where used in this research. It would be interesting to change the groups, so not that we are only looking at different professional careers. But look at people with different economic backgrounds or between countries. Looking to different groups can also mean that we can see if people from the same group are connected by having the same motivation or goal in life and that we can see this back in their language? And if we use the same communication style does that mean that we then also have a similar personality?

Crystal is using a disclaimer on their website telling us that profiles are a predicted personality based on social media information and other public data. Giving us a "best guess" about a persons preferred communication style. This research used the data of Crystal to later analyse it with LIWC, we should be aware that the conclusions we made about a certain group are based on a "best

guess" method. And that not all the persons in a certain group communicated in the same way. For this research the data set contain only people who have a LinkedIn account. Without LinkedIn account Crystal was not able to predict a personality mostly because we got no results at all from Crystal or the certainty was below 10%. An other limitation is that we don't know exactly how Crystal is getting the information from the online profiles of a person.

This research is showing not only that we can predict how people would like to communicated but also that the language is telling us much more! And that we can use communication to influence people. Further research in languages can give us more interesting outcomes. This time we looked at personality traits and communication. Other research already has shown that language and health are connected with each other. Further research can pick up where we left. In the end we should find out if language is something like our DNA which can be tracked back to one person, one personality and is unique for every person around us. Finding out if languages is a fingerprint of a person would be the next step in this field of research.

To sum up, my final recommendations for future research follows from the limitations of this research.

1) Change the groups from career perspective to different groups like income, different ages, nationality, different communities.

2) Using Crystal without LinkedIn profile, at this point Crystal was not able to give us outcomes when the person was missing a LinkedIn account, but it can be very interesting to do this again when other profiles give enough information to Crystal to predict their DISC profile.

Reference

ALLC. (2013). Literary and Linguistic Computing Advance Access. Oxford University Press.

Crystal. (2016, October). Understanding DISC. Retrieved from Crystalknows: https://www.crystalknows.com/disc

Crystalknows. (2016, September). Crystal. Retrieved from © Crystal Project Inc.: https://www.crystalknows.com

Davison, K. P., Pennebaker, J. W., & Dickerson, S. S. (2000). Wo Talks? The Social Psychology of Illness Support Groups. American Psychologis.

Frost, A. (2016, August 11). Blog: Here's What Happened When I Let an App Dictate How I Wrote Emails. Retrieved from redbooth: https://redbooth.com/blog/crystal-test-drive?utm_content=buffer5d8a6&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer

Homes, D. I. (1998). The Evolution of Stylometry in Humanities. The College of New Jersey, Vol. 13, No. 3.

Hunt, E. (2015, May 19). Crystal knows best.. or too much? The diconcerting new email advice service. (The Guardian) Retrieved from The Guardian: https://www.theguardian.com/media/2015/may/19/crystal-knows-best-or-too-much-the-disconcerting-new-email-advice-service

Ireland, M. E., & Slatcher, R. B. (2011). Language Style Matching Predicts

Relationship Initiation and Stability. Psychological Science, 22(1), 39-44.

Kloosterman, F. (2016, February 29). Personal Branding to the max, want Crystalknows 'kent' je. Retrieved from pintapeople: http://pintapeople.com/personal-branding-to-the-max-want-crystalknows-kent-je/

Krueger, D. (1998). Personality characteristics of the small business entrepreneur. Journal of Business and Entrepreneurship.

Krueger, D. (2000). Characteristics of the female entrepreneur. Journal of Business and Entrepreneurship.

Labrecque, L. I. (2011). Online personal branding: processes, challenges, and implications. Journal of Interactive Marketing(25(1)), 37-50.

Linoff, G. S. (2011). Data mining techniques: for marketing, sales, and customer relationship management. John Wiley & Sons.

Ng, S. H. (1993). Power in language: Verbal communication and social influence. Sage Publications, Inc.

Pennebaker. (2015). How it works. Retrieved from LIWC: http://liwc.wpengine.com/how-it-works/

Pennebaker, J. W. (1991). Linguistic styles: language use as an individual difference. Journal of personality and social psychology, 77(6), 1296.

Pennebaker, J. W., & Graybeal, A. (2001). Patterns of Natural Language

use: Disclosure, Personality, and Social Integration. Saga journals, Volume 10, No. 3.

Pennebaker, J. W., & King, L. A. (1999). Linguistic Styles: Language Use as an Individual Difference. Journal of Personality and Social Psychology, No.6, 1296-1312.

Pennebaker, J. W., Mehl, M. R., & Niederhoff, K. G. (2003, February). Psychological aspects of natural language use: Our words, Our selves. Annaual Review of Psychology, Vol. 54, 547-577.

Pratt, S. (2015, June 25). Crystal. Retrieved from Socialtalent: http://www.socialtalent.co/blog/crystal-for-linkedin

Ramsey, R. P. (1997). Listening to your customers: The impact of perceived salesperson listening behavior on relationship outcomes. Journal of the Academy of marketing Science, 25(2), 127-137.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and Computerized Text Analysis Methods. Journal of Language and Social Psychology.

Watson, D., & Klohnen, E. (2004). Match makers and deal breakers: Analyses of assortative mating in newlywed couples. Journal of Personality, 71, 1029-1068.

Williams, O. (2015, April 15). Creep on your friends. Retrieved from The

next web: http://thenextweb.com/apps/2015/04/15/creep-on-your-friends/

Exhibit 1. Overview Crystal

The overview is from Prof. Willem Verbeke with a DISC profile Di. This tells us that his primary type is <u>D</u>ominant followed by <u>I</u>nfluencing.

Prof. wants to get to know their colleagues so keeping the conversation strictly professional may bore them. **Don'**t take yourself too seriously and they will do the same.

When **speaking** to Prof....

| Get right to the bottom line | 👍 | 👎 | Stick to the big picture | 👍 | 👎 |

| Project boldness and confidence (even at the risk of sounding impolite) | 👍 | 👎 |

| Expect them to interrupt | 👍 | 👎 |

When **emailing** Prof....

| State your purpose for the email in the first sentence | 👍 | 👎 |

| Write 3 sentences or less | 👍 | 👎 |

| Write with short casual language and abbreviations | 👍 | 👎 |

| Use data to prove a point | 👍 | 👎 |

When **working with** Prof....

| Always focus on the team's overall goal | 👍 | 👎 |

| Offer blunt constructive criticism | 👍 | 👎 |

| Confront conflict in person rather than via email | 👍 | 👎 |

| Spend time exploring new ideas | 👍 | 👎 |

It **comes naturally** to Prof. to...

Feel comfortable making decisions with limited information 👍 👎

Complete projects more quickly than others 👍 👎

Prefer to improve an existing document rather than write a new one 👍 👎

Immediately project assertiveness and confidence 👍 👎

Exhibit 2. Positive emotions, negative emotions and sad between the groups.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
posemo	Between Groups	81,573	5	16,315	1,838	,106
	Within Groups	2255,137	254	8,878		
	Total	2336,710	259			
negemo	Between Groups	28,526	5	5,705	3,103	,010
	Within Groups	467,031	254	1,839		
	Total	495,557	259			
sad	Between Groups	1,984	5	,397	3,442	,005
	Within Groups	29,281	254	,115		
	Total	31,265	259			

Exhibit 3. Insights, causation, certainty and differentiation between and within groups.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
insight	Between Groups	129,638	5	25,928	4,917	,000
	Within Groups	1339,385	254	5,273		
	Total	1469,023	259			
cause	Between Groups	65,819	5	13,164	12,238	,000
	Within Groups	273,218	254	1,076		
	Total	339,037	259			
certain	Between Groups	79,964	5	15,993	7,936	,000
	Within Groups	511,848	254	2,015		
	Total	591,811	259			
differ	Between Groups	84,824	5	16,965	10,337	,000
	Within Groups	416,840	254	1,641		
	Total	501,664	259			

Exhibit 4. Affiliation, reward and risk between and within groups.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
affiliation	Between Groups	29,616	5	5,923	4,691	,000
	Within Groups	320,744	254	1,263		
	Total	350,359	259			
reward	Between Groups	53,613	5	10,723	7,900	,000
	Within Groups	344,746	254	1,357		
	Total	398,359	259			
risk	Between Groups	9,835	5	1,967	2,074	,069
	Within Groups	240,913	254	,948		
	Total	250,747	259			

Exhibit 5. Work and leisure between and within groups

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
work	Between Groups	129,538	5	25,908	9,618	,000
	Within Groups	684,195	254	2,694		
	Total	813,733	259			
leisure	Between Groups	61,213	5	12,243	12,172	,000
	Within Groups	255,473	254	1,006		
	Total	316,685	259			

Exhibit 6 – LIWC2015 Output Variable Information

The table below provides a comprehensive list of the default LIWC2015 dictionary categories, scales, sample scale words, and relevant scale word counts.

Category	Abbrev	Examples	Words in category	Internal Consistency (Uncorrected α)	Internal Consistency (Corrected α)
Word count	WC	-	-	-	-
Summary Language Variables					
Analytical thinking	Analytic	-	-	-	-
Clout	Clout	-	-	-	-
Authentic	Authentic	-	-	-	-
Emotional tone	Tone	-	-	-	-
Words/sentence	WPS	-	-	-	-
Words > 6 letters	Sixltr	-	-	-	-
Dictionary words	Dic	-	-	-	-
Linguistic Dimensions					
Total function words	funct	it, to, no, very	491	.05	.24
Total pronouns	pronoun	I, them, itself	153	.25	.67
Personal pronouns	ppron	I, them, her	93	.20	.61
1st pers singular	i	I, me, mine	24	.41	.81
1st pers plural	we	we, us, our	12	.43	.82
2nd person	you	you, your, thou	30	.28	.70
3rd pers singular	shehe	she, her, him	17	.49	.85
3rd pers plural	they	they, their, they'd	11	.37	.78
Impersonal pronouns	ipron	it, it's, those	59	.28	.71
Articles	article	a, an, the	3	.05	.23
Prepositions	prep	to, with, above	74	.04	.18
Auxiliary verbs	auxverb	am, will, have	141	.16	.54
Common Adverbs	adverb	very, really	140	.43	.82
Conjunctions	conj	and, but, whereas	43	.14	.50
Negations	negate	no, not, never	62	.29	.71
Other Grammar					
Common verbs	verb	eat, come, carry	1000	.05	.23
Common adjectives	adj	free, happy, long	764	.04	.19
Comparisons	compare	greater, best, after	317	.08	.35
Interrogatives	interrog	how, when, what	48	.18	.57
Numbers	number	second, thousand	36	.45	.83
Quantifiers	quant	few, many, much	77	.23	.64
Psychological Processes					
Affective processes	affect	happy, cried	1393	.18	.57
Positive emotion	posemo	love, nice, sweet	620	.23	.64
Negative emotion	negemo	hurt, ugly, nasty	744	.17	.55
Anxiety	anx	worried, fearful	116	.31	.73
Anger	anger	hate, kill, annoyed	230	.16	.53
Sadness	sad	crying, grief, sad	136	.28	.70
Social processes	social	mate, talk, they	756	.51	.86
Family	family	daughter, dad, aunt	118	.55	.88

Category	Abbrev	Examples	Words in category	Internal Consistency (Uncorrected α)	Internal Consistency (Corrected α)
Friends	friend	buddy, neighbor	95	.20	.60
Female references	female	girl, her, mom	124	.53	.87
Male references	male	boy, his, dad	116	.52	.87
Cognitive processes	cogproc	cause, know, ought	797	.65	.92
Insight	insight	think, know	259	.47	.84
Causation	cause	because, effect	135	.26	.67
Discrepancy	discrep	should, would	83	.34	.76
Tentative	tentat	maybe, perhaps	178	.44	.83
Certainty	certain	always, never	113	.31	.73
Differentiation	differ	hasn't, but, else	81	.38	.78
Perceptual processes	percept	look, heard, feeling	436	.17	.55
See	see	view, saw, seen	126	.46	.84
Hear	hear	listen, hearing	93	.27	.69
Feel	feel	feels, touch	128	.24	.65
Biological processes	bio	eat, blood, pain	748	.29	.71
Body	body	cheek, hands, spit	215	.52	.87
Health	health	clinic, flu, pill	294	.09	.37
Sexual	sexual	horny, love, incest	131	.37	.78
Ingestion	ingest	dish, eat, pizza	184	.67	.92
Drives	drives		1103	.39	.80
Affiliation	affiliation	ally, friend, social	248	.40	.80
Achievement	achieve	win, success, better	213	.41	.81
Power	power	superior, bully	518	.35	.76
Reward	reward	take, prize, benefit	120	.27	.69
Risk	risk	danger, doubt	103	.26	.68
Time orientations	TimeOrient				
Past focus	focuspast	ago, did, talked	341	.23	.64
Present focus	focuspresent	today, is, now	424	.24	.66
Future focus	focusfuture	may, will, soon	97	.26	.68
Relativity	relativ	area, bend, exit	974	.50	.86
Motion	motion	arrive, car, go	325	.36	.77
Space	space	down, in, thin	360	.45	.83
Time	time	end, until, season	310	.39	.79
Personal concerns					
Work	work	job, majors, xerox	444	.69	.93
Leisure	leisure	cook, chat, movie	296	.50	.86
Home	home	kitchen, landlord	100	.46	.83
Money	money	audit, cash, owe	226	.60	.90
Religion	relig	altar, church	174	.64	.91
Death	death	bury, coffin, kill	74	.39	.79
Informal language	informal		380	.46	.84
Swear words	swear	fuck, damn, shit	131	.45	.83
Netspeak	netspeak	btw, lol, thx	209	.42	.82
Assent	assent	agree, OK, yes	36	.10	.39
Nonfluencies	nonflu	er, hm, umm	19	.27	.69
Fillers	filler	Imean, youknow	14	.06	.27

"Words in category" refers to the number of different dictionary words and stems that make up the variable category. All alphas were computed on a sample of ~181,000 text files from several of our language corpora (see Table 2). Uncorrected internal consistency alphas are based on Cronbach estimates; corrected alphas are based on Spearman Brown. See the Reliability and Validity section below. Note that the LIWC2015 dictionary generally arranges categories hierarchically. There are some exceptions to the hierarchy rules. For example, *Social processes* include a large group of words that denote social processes, including all non-first-person-singular personal pronouns as well as verbs that suggest human interaction (talking, sharing) -- many of these words do not belong to any of the *Social processes* subcategories. Another example is *Relativity*, which includes a large number of words that cannot be found in any of its subcategories.