Computers in Human Behavior 59 (2016) 39–48

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

The sociability score: App-based social profiling from a healthcare perspective

Paul Eskes [a], Marco Spruit [a], Sjaak Brinkkemper [a], Jacob Vorstman [b,1], Martien J. Kas [c,*,1]

[a] *Department of Information and Computing Science, Utrecht University, Princetonplein 5, Utrecht, The Netherlands*
[b] *Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Heidelberglaan 100, Utrecht, The Netherlands*
[c] *Department of Translational Neuroscience, Brain Center Rudolf Magnus, University Medical Center Utrecht, Universiteitsweg 100, Utrecht, The Netherlands*

ARTICLE INFO

Article history:
Received 23 July 2015
Received in revised form
4 January 2016
Accepted 18 January 2016
Available online xxx

Keywords:
Data mining
Psychological health
Smartphone use
Sociability
Social deficits
Social profiling

ABSTRACT

As the smartphone becomes an integral part of our lives, its value as a rich data source reaches an increasing potential. Several previous studies have used smartphone-derived data to discover relationships between user characteristics and different types of smartphone use. However, none tried to use smartphone data to capture an individual's social behavior into one profile, aimed at providing additional information for the diagnostic evaluation of social deficits. This study presents a novel way of combining different modalities of smartphone data for the creation of sociability profiles using a scoring mechanism that allows for easy addition and removal of data sources. Following installation of the smartphone application, data is being sampled in the background to allow for the assessment of spontaneous smartphone use. Sociability scores were based on the integration of social communication and social exploration scores derived from smartphone use and environmental data sampling (e.g., GPS and external Bluetooth signals). Finally, we have applied our Sociability model to create social profiles of ten test subjects as a baseline for future studies. This pilot study provided insight in the usability of the individual sociability scores for future smartphone application to provide longitudinal objective measures of normal and atypical human social behavioral profiles in their natural environment.

1. Introduction: smartphone-based social profiling in healthcare

According to a survey by Ernst and Young, Dutch citizens are dissatisfied about IT-innovation in the healthcare sector, which is explained by the fact that healthcare specialists provide human-centered services and are not focused, nor trained, on applying IT-solutions to improve their business processes (Ernst & Young, 2011). This observation also suggests a wealth of opportunities present in this sector for new technologies, one of which consists of the exploration of ways smartphones can aid in professional healthcare. The larger the role of smartphones becomes in our lives, the more interesting these devices will become for healthcare, given that the information held by these smartphones could provide objective insights into the owner's lifestyle and possibly, into their psychological wellbeing. Currently, many data mining techniques exist that can extract data from smartphones about the user's smartphone use. But in order to analyze aspects of social behavior based on a large set of extracted data from a large group, a validated, scientific model should be developed. Subsequently, this information can be used in the psychological domain to create distinctive social profiles and thus create valuable insights in a person's level of sociability. Finally, this information can be a valuable addition in a clinical context, with the potential to contribute to the accuracy of medical diagnoses in the cognitive-behavioral domains and therefore improving the overall efficiency of subsequent treatment.

Currently, the department of Translational Neuroscience and the department of Psychiatry at the University Medical Center Utrecht, in cooperation with a third party software vendor, develop a smartphone application with the aim to create additional digital assistance for the objective and longitudinal observation of social behavior and, potentially for the diagnosis and early detection of patients with (possible) social deficits. Furthermore, in view of treatment efficacy monitoring of these patients, there is a great need to obtain information regarding spontaneous human social

* Corresponding author.
E-mail addresses: J.A.S.Vorstman@umcutrecht.nl (J. Vorstman), m.j.h.kas@umcutrecht.nl (M.J. Kas).
[1] Equal contributors.

http://dx.doi.org/10.1016/j.chb.2016.01.024
0747-5632/

behavior in their natural environment. Current methods are mainly based on self-reports, obtained either through questionnaires, or through life or phone-based interviews. The two largest disadvantages of these currently used methods are the restricted amount of information these sessions deliver and the questionable reliability of the information given the inherent subjective quality of the data. As literature states, self-reported statements can only be interpreted when handled with great care, as people may change the truth consciously or unconsciously to get a desired outcome or because they have a wrong impression of their own situation (Straka, Fish, Benson, & Suh, 1997; De Reuver et al., 2012).

The smartphone as an objective instrument eliminates both of these disadvantages that are connected with the current diagnostic method, since a smartphone can collect information both extensively and objectively. Therefore it is interesting to examine the role smartphones can play in current diagnosing methods for social deficits. Many studies show the possibilities of smartphones as a data source for all types of social purposes such as user profiling, user tracing, and activity recognition. But little research has been done to utilize these possibilities to fit healthcare purposes (e.g. Meulendijk, Meulendijks, Jansen, Numans, & Spruit, 2014), and more specifically in the psychological healthcare domain. To discover the potential of smartphones for clinical purposes, additional research is required that may uncover the possibilities of smartphone data in the diagnosis and treatment of people with possible social deficits. As an introduction to this topic, this research involves uncovering the potential of the smartphone as a measurement instrument for the psychological healthcare domain. We therefore formulate the following research question.

Can individual social profiles be created for psychological healthcare purposes based on smartphone usage and smartphone-registered behavior?

The first step in answering the main question is to identify and define the different factors that can be considered as the building blocks of a social exploration profile. In the context of this research, all of these factors derived from smartphone-retrieved data, including smartphone-activity data and data retrieved from smartphone sensors, which are used to directly or indirectly describe an aspect of a user's sociability. It should be taken into consideration however that some of these factors may be explained by certain user characteristics and demographics and are of no direct consequence regarding a person's sociability.

This research has relevance for multiple target groups when looked at from different perspectives. First, the scientific community gains new insights into the possibilities of smartphone data for social studies and, more specifically, into using information about an individual's smartphone use as an additional source for describing the sociability of an individual. From a business perspective, a description about an individual's sociability levels can function as an additional source for physicians during the process of diagnosis, which ideally can gain hospitals an increased effectiveness and efficiency of several treatments in the psychological domain, reducing treatment times, waiting lists and overall treatment costs. Finally, from a social point of view, patients and hospitals both benefit as it is also in their best interest to improve quality of diagnosis and treatment.

The structure of this paper is as follows. First, a theoretical background will be provided for this topic in the form of a systematic literature review; exploring the current best practices in the field of smartphone mining and sociability. Then, both the research and data mining method will be described in the research approach chapter to provide structure to the research process. Next, the sociability model will be presented, which will be tested subsequently by applying the model in a test group of 10 individuals of which the results will be presented in the results chapter. Finally, in the conclusion and discussion sections we evaluate the results and describe future research opportunities.

2. Theoretical background

First, we select the top predictors that may serve as confounding factors for the creation of a social profile. As the theory of behavior model suggests, we should determine the possible impact several categories have on our research, to avoid having confounding factors determining the social profile. The categories described by the theory of behavior include consumer attributes, user-context, service characteristics, intentions, and technology enablers (Maheshwaree, Ylä-Jääski, Verkasalo, & Hämmäinen, 2009). When observing people's social activities, the biggest impact is created by the person confounding factors, which include user characteristics and user demographics (Steg, Buunk, & Rothengatter, 2008). These factors are user-specific, and are proven to influence social behavior in several ways. The social cognitive theory confirms that personal factors influence social behavior and adds to the importance of environmental factors which partly overlap with the user-context attributes described in the theory of behavior model (Bandura, 1986). The remaining categories intentions, service characteristics and technology enablers only indirectly influence social behavior through increased/reduced smartphone use and therefore are not confounding factors for a sociability profile. As an example of the category intentions, performance expectancy influences the consideration to use the smartphone for communication, but when the user decides not to use the smartphone for communication, he will choose a different medium for communication which does not make him less sociable. Also, having unlimited internet access or a smartphone with a high-capacity battery will increase overall smartphone use which will also result in increased social smartphone use. This however does not implicate that the person is more sociable, while a sociable person will again search for alternative ways of communication when the phone battery is fully drained or when internet access is unavailable. In conclusion, only the personal factors and user-context factors are influences that should be accounted for when observing social smartphone use in more detail. We will explain the explicit factors in more detail in Sections 4 and 5.

The first data source category is social media usage, including social behavior on platforms such as Facebook, Twitter, and Tagged. In the context of social media, research to date is limited to studies that describe social media usage (e.g. Buijs & Spruit, 2015), regardless whether the platform is either a PC, smartphone or tablet. Considering again the Theory of Behavior model by Maheshwaree et al. (2009), the findings associate personal characteristics with social media usage, making personal factors again a potential confounding factor, in this case, for examining social media usage. No research however has been done to investigate the influence of user-context on social media usage, making it unclear whether user-context has the same influence on social media usage as it has on smartphone use in general. Additionally, factors that could be placed under the remaining categories of the theory of behavior model; intentions, service characteristics and technology enablers, are not found in present literature. As for the largest difference between general social media usage for the PC and the smartphone, Kaikkonen concludes that people use mobile internet mostly for following social media sites. Desktop computers on the other hand are used for active contribution to social websites, which was less common when using mobile devices (Kaikkonen, 2008). Twitter usage was found to be used more extensively on the smartphone, which was explained due to the short attention span required for tweeting (Grace, Zhao, & Boyd, 2010). In short, our literature review indicates that several research gaps exist concerning social media usage on smartphones, which underlines the possibilities for the data that are being collected for this study.

The second data source category is the use of Bluetooth signals for sensing social situations. Two applications of Bluetooth signals were found to determine the type of social situations users were involved in. Nicolai and Kenn (2006) chose to identify activities afterward during interventions with the users, based on the identification of other smartphones and the total amount of signals in the surroundings. Yan, Yang, and Tapia (2013) used the same data sources, but applied several computational techniques to distinguish recurrent daily activities and to additionally create entropy maps visualizing the Bluetooth density of a user's environment during the week. Both methods use a combination of Bluetooth signal identification and the quantity of Bluetooth signals in the surroundings as a base for their analysis but both in different manners, underlining the possibilities offered by Bluetooth signals for future studies.

The last topic concerning sociability is the information we can derive from the localization of smartphone users, for which a combination of data sources is required that connects physical locations to symbolic locations or vice versa (Hightower & Boriello, 2001). Several examples exist in literature that achieve localization in distinctive ways, of which LifeMap can be considered the most promising one. LifeMap combines five different data sources leading to the identification of POIs with 91% accuracy within an error bound of 25,6 m. As derived from best practices, it is common to use GPS as a base and add other data sources to fill data gaps caused by poor indoor GPS reception and to apply symbolic meaning to the physical coordinates of a GPS signal. In extension of the localization process, although not one hundred percent accurate, it is possible to create movement paths onto street maps by an algorithm that makes use of predictive modeling. Such methods have the disadvantage of frequent location determination and the high battery consumption resulting from this high retrieval rate. Still, Bierlaire, Chen, and Newman (2010) in their research perceive such a method to be well suited for the sparse and sometimes inaccurate data delivered by the smartphone's GPS utility. Further research should investigate this viewpoint by Bierlaire, Chen and Newman, while little research exists for smartphone-based localization. Besides movement patterns as an extension of localization, Adams, Phung, and Venkatesh (2008) extend the concept of localization by enriching locations with information about people in the environment, time and duration to discover so called social rhythms; recurrent activities that can be characterized by the place, the duration, the time of occurrence and other people present during the activity. In the context of sociability and healthcare, researchers can use anomalies in social rhythms to identify social withdrawal or perhaps even mental illness exacerbation.

3. Research approach

The data collection tool used within this research to abstract data describing regular smartphone usage is called BeHapp, a tailor-made smartphone application for the Android operating system developed by a Dutch software development company. The application collects data from multiple data sources, from which we selected the following to be valid and reliable enough for further data analysis: incoming/outgoing calls, missed calls, call duration, application accessing, localization of the user with GPS signals and the counting of Bluetooth signals in the area. In addition to each of these events, the time the event occurred and the location of the user, at the time the event occurred, is registered as an attribute.

For the overall data mining process, we chose CRISP-DM as the main guideline for the knowledge discovery process in this research (Chapman et al., 2000; Shearer, 2000). Following the phases of CRISP-DM, the first step is to create a domain understanding by defining a data mining problem and by creating a preliminary plan on how to achieve the objectives. This was done by interviewing the two experts in the fields of sociology and psychiatry, who initiated this project, to discuss the requirements and outcomes of the method. The second step is to build a general understanding of the data, for which interviewing is required with the developers of the application. This phase also includes the acceptance testing for verifying the validity and reliability of the data generated by the application. The testing will most likely lead to a new development phase, where bugs are fixed and additional features are implemented. Simultaneously, interviews with domain experts will be conducted to define the goals and required variables that are related to the social profile of the smartphone owner. Next, the collected data will be prepared, modeled and evaluated to get an impression of the final social profile and to make alterations to the previous definition of this profile if needed. When both the smartphone application is technically accepted and the method for defining a person's social profile is found acceptable, the research can continue with a second iteration of the CRISP-DM process. This time, the method will be used to create social profiles of 10 individuals, of which data will be collected over 1,5 week. Based on the results, the experts will provide their opinions on the satisfactory level that the method provides for health care professionals as a part of the last evaluation phase.

For the processing and analysis of the data in preparation of determining the sociability score, several different tools were used, in line with the Usage Mining Method of Pachidi, Spruit, Van der Weerd, (2014). First, all data collected by the BeHapp application was sent to and stored in a MySQL database with access through phpMiniAdmin. From the web application, csv-files were exported, which functioned as exchange files between the different analysis tools. Then, Rapidminer Studio Free Edition was used as our data analytics application to cleanse, filter and transform the data and eventually analyze the data. For the creation of the final data set, the data had to undergo the following processes: removal of double entries, removal of empty entries, removal of unanswered outgoing phone calls, filtering of data outside the time period, filtering of unreliable test subjects and filtering of faulty data. Additionally, SPSS was used for statistical analysis purposes and to provide visualizations that could not have been created by the standard toolset of Rapidminer Studio. An operational version of the sociability scoring method was eventually created in Rapidminer, to provide experts with a tool to score new test subjects. The only effort to be made for scoring new individuals is the replacement of the old CSV-file with a new CSV-file exported from the MySQL database.

4. Development of the sociability score

As found during the literature study, the definition of the term ‘sociability’ that we will maintain during this study will be:

“The tendency to affiliate with others and to prefer being with others to remaining alone.” (Cheek & Buss, 1981).

This definition consists of two parts; ‘the tendency to affiliate with others’, which expresses itself in social acts that contribute to certain social situations. And ‘the tendency to prefer being with others to remaining alone’, which can be explained by the crowdedness of situations an individual puts himself into. As we derive from this definition, utterances of sociability can be divided into social acts and social exploration, of which the terms social acts and communication will be used interchangeably. In this chapter we will describe the concept of sociability in more detail, combining information from the literature, expert interviews and the data understanding phase to create a sociability model that is designed to assign a sociability score to a smartphone user based on his smartphone data.

4.1. Domain understanding

During the expert interviews, the experts concluded that to describe a user's social acts you can identify three metrics; frequency of social acts, duration of the interaction and diversity of communication partners. Duration was chosen as a metric, while one expert mentioned that in his experience, schizophrenic patients have a shorter length of utterance (the length of an average sentence), which hypothetically will have an effect on the eventual conversation duration.

In line with the definition part ‘to prefer being with others to remaining alone’, the experts added the social environment as a measure of the amount of people present during user's daily situations. The term social environment later transformed into social exploration, which includes social density, but also additional dimensions derived from the expert interviews, like the movement range, the variation in places visited, the duration of place visits and the diversification in movement patterns.

The movement range was added under the assumption that people with social deficits have a limited movement range; they feel less compliant traveling large distances to attend a certain social situation. Also, variation in places visited is added to capture the diversity in places a person normally visits during a certain time span. Furthermore, one of the experts mentioned duration to uncover the subjects' attitude towards the presence of other people for an extended period of time, from which the duration of place visits was derived. Finally, movement patterns also became part of social exploration, while people with social deficits like schizophrenia tend to avoid crowded places, and therefore alter their routes based on the crowdedness of a certain situation.

4.2. Data understanding

Concerning communication, only data describing phone call behavior has been found reliable enough to add as a data source for the sociability model. From this data source, all aforementioned dimensions frequency, diversity and duration can be used as factors describing overall call behavior. A distinction will be made here in calls that are incoming and calls that are outgoing to distinguish between active senders and active receivers. The experts make specific statements about the role of the user in a conversation: being either a sender or a receiver.

Localization of users can be approached by determining either the physical or the symbolic location of the user and subsequently linking the other (manually or automatically) to the one observed (Hightower & Boriello, 2001). For the sociability model we decided not to apply physical localization, because creating a score out of physical places would require additional knowledge about these particular places, which would make the scoring unnecessarily complex. For the scoring method we chose to focus only on symbolic locations which are formed by the variables social density, duration and the distance from home. The biggest advantage of these variables is that the numbers can be compared objectively regardless of what the exact physical location is. A consequence of this method is that the factor ‘variation in places visited’ becomes redundant, as the variation in symbolic locations can be described by comparing the values of the associated attributes. As duration appeared to be an unreliable variable, an alternative way had to be found to determine symbolic locations and to give value to the social density and distance. We did this by assigning a score to both variables separately as we will explain in more detail in the following section. In the context of movement mining, it appears too premature to sort out a user's movement paths or (recurrent) physical activities for now, mainly because of the complexity caused by the low position registration rate.

4.3. The sociability model

The conceptual version of the sociability model is presented in Fig. 1. It shows the different dimensions of sociability and the variables derived from the expert interviews that contribute to these dimensions. The model has been designed under the assumption that every data source is available. However, in the context of BeHapp, several parts of the model could not be tested, due to technical restrictions of the application. For this reason, an overlay is created in Fig. 1 filtering the variables disapproved based on their validity or reliability. We will describe the representation of the variables in the model in more detail below.

4.3.1. Variables scores

Each of the variable scores is calculated in a similar manner, differing only in nuances. The distinctive scores can be calculated with either a social act, travel distance or a Bluetooth density formula. The social act formula in Eq. (1) first calculates the standard score for the particular user and a specific event X, using the average μ and standard deviation σ of the total population for that same event X. This social act score can be explained as the amount of standard deviations the particular user deviates from the average in a normal distribution.

In order to avoid a negative score, all scores have to be shifted to a positive domain by adding a factor i to the scores. Note, however, that in a normal distribution, it is only possible to shift all scores to a positive domain if the distribution is shifted with $i=\infty$ while a Gaussian curve includes all possible scores within the domain of ∞ to ∞ Therefore, boundaries have to be set for what are expected to be possible scores per research. Another reason to limit the domain is that the wider the score range chosen, the more the scores will cluster together, making it harder to separate subtle differences between test subjects. Therefore, the factor i should be chosen considerately to avoid the scores from either clustering or falling outside the score range. The first consideration was to choose the six sigma strategy, a term frequently used in business performance management, which prescribes to use a maximum of 6 standard deviations to maintain an efficiency of 99,99966% (or 3,4 errors per million scores) (Barney & McCarty, 2003). However, in case of this research, clustering of scores occurred under this strategy, which led to the decision to shift the scores by adding 3 to each standard score, as in 99,73% of the cases the standard score deviates from −3 to +3. In this case, there still exists a small chance that 1 person out of 370 falls outside of this scale, causing an error, but we assume this chance to be too small in a sample set of 10 subjects. After the score calculations, this assumption was verified, as the test subject scores deviate from a minimum of 15,93 to a maximum of 95,82. In case the same formula will be used for a bigger sample set, one might consider increasing the amount of standard deviations, making again the tradeoff between the subtlety in score differences and the probability of errors to find the optimal balance for that particular situation. To come to a score eventually, we divide the shifted standard score by the maximum standard score, which equals two times i (i.e. two times the amount of standard deviations used for the shift). Which would be six for this research as $i = 3$. Finally, the score is multiplied by a hundred to create a score between 1 and 100.

The social act formula for a particular user and a specific event X.

$$social\ act\ score = \left(\left(\left(\frac{X-\mu}{\sigma}\right)+i\right)\Big/2i\right)*100 \tag{1}$$

Range operationalization of the social act formula with 3-sigma limits.

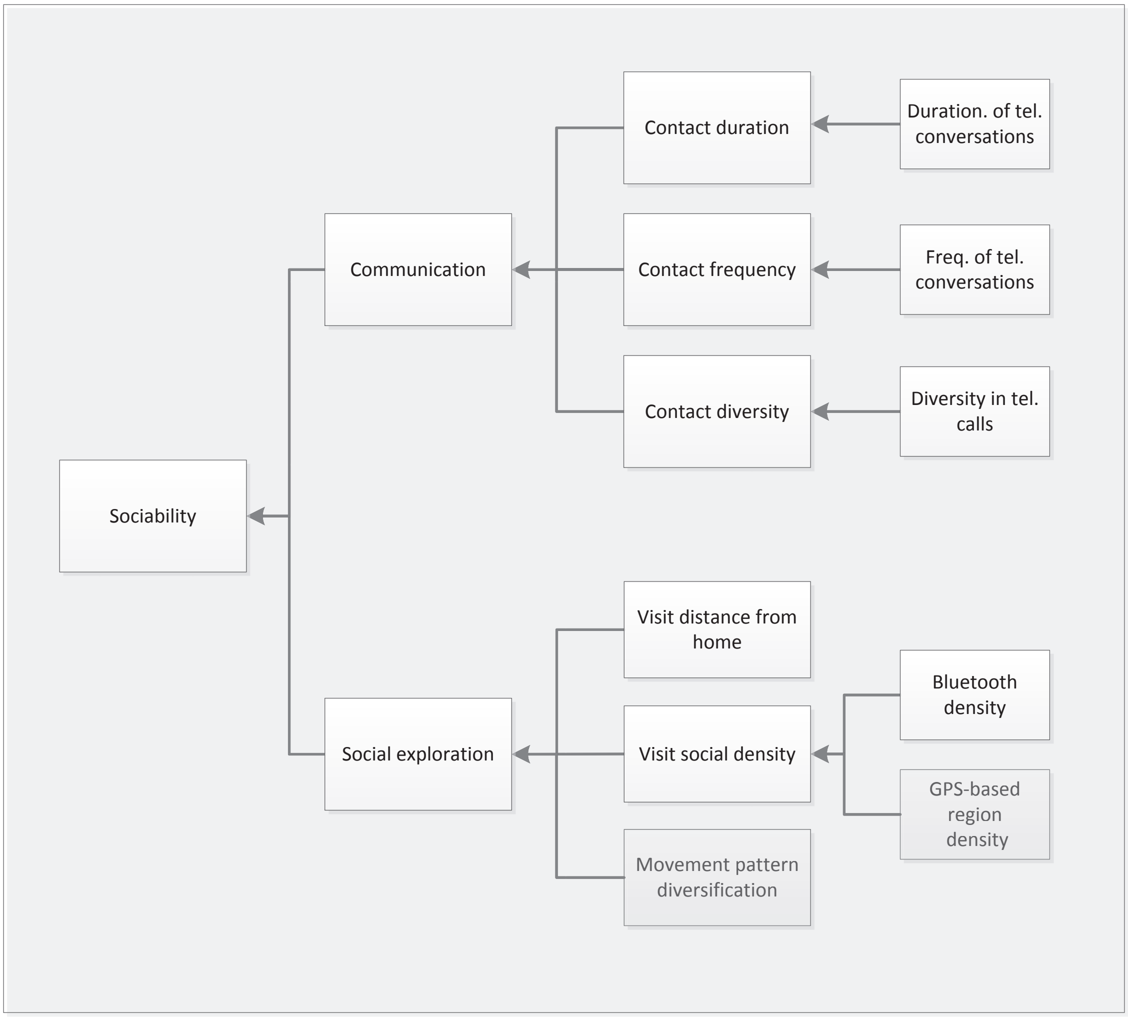

Fig. 1. The Sociability model (including hypothesized factors in gray).

$$\text{social act score} = \left(\left(\left(\frac{X-\mu}{\sigma}\right)+3\right)\Big/6\right)*100 \quad (2)$$

Besides the scoring of social act frequencies, the average call duration is also transformed using the same formula, where in this case the value of X is equal to the average conversation duration in seconds over all either incoming or outgoing telephone calls.

To give a score to the distance traveled by a user, first for all of the GPS-positions the distance had to be calculated to the subject's home. The subjects had to fill in the GPS-coordinates of their home before participating in the study. Then, we use a clustering method on the available data to group the data per hour and per day respectively, and taking the average distance for each of these steps. As a threshold we chose for a minimum of 3 data points before grouping, meaning that if for instance the average over an hour was derived from only 2 data points, the whole hour was omitted from further grouping. Also, if a subject only has effective data from 2 days in total, the person did not get a distance score assigned. The total average distance of a subject is subsequently transformed into a score using the same formula used for social act scores as shown in Eq. (1). For the Bluetooth density score, a similar method is used as for the travel distance score, but instead of taking the distance average, the average Bluetooth signal count is taken. The clustering of the data occurs under the same circumstances just as the scoring of the average amount of Bluetooth signals.

4.4. Confounding factors

As we recall from the systematic literature review, the personal factors and user-context factors are influences that should be accounted for when observing social smartphone use in more detail in the context of measuring sociability (Bandura, 1986; Maheshwaree et al., 2009; Steg et al., 2008). Personal factors include gender, age, profession, culture etcetera, but these factors can be controlled by picking the right sample set. A personal factor which cannot be observed directly is personality, which consists of several different dimensions that should be tested separately to create an extra personality profile. For both the preservation of validity and the additional exploratory reasons, a personality profile will form an essential part of the eventual sociability profile. This profile will show to what extent the scores can be explained by an individual's personality by using percentages. Furthermore, user-context factors also play a role in overall smartphone use, while for example the presence of WiFi-signals, the smartphone battery capacity and the smartphone specifications may all indirectly play a role in the execution of social acts (Maheshwaree et al., 2009). Under the assumption that a linear relationship exists between smartphone use in general and social smartphone use if not influenced by personal or user-context factors, a profile of general smartphone use is required to put the social smartphone use into context. For instance, a low sociability score can in this way be explained when the subject is not a frequent smartphone user. This can be the case when the subject does most of his electronic communication using the telephone at work or at home. Concluding, additional reference points are required in the form of a smartphone use profile and a personality profile to include the confounding factors and to put the results of the sociability profile into context. The personality profile will be created by filling in a questionnaire. The smartphone use profile will only consist of the Application Activity (AA) frequency score, while the AA duration appeared not reliable enough to draw conclusions upon and the AA categorization (diversity) missed data due to an unfortunate bug in the latest version of BeHapp. The AA frequency score will be determined using the same method to describe the frequency of CIs and COs.

4.5. The sociability score

The last step is choosing a way of representing the sociability of a certain user in quantitative figures. However, the largest problem to cope with when creating scores is the fact that the application does not register every event properly and may miss a phone call or SMS message, as explained in the reliability section. For this reason, it is unreliable to conclude something about a person's sociability, while data may be missing. For instance, the database could contain 10 phone calls for a certain subject, of which could be concluded that the subject is an average caller, while literature states the average amount of phone calls per subject is 8 phone calls a week. In real life however, this person could have called way more frequently, but not having these extra calls registered, making this person in real life a more frequent caller. A possible solution for this issue could be to use prediction models, using manually gathered phone call data as a training set. However, we do not have a test group which is large enough to provide enough data for reliable prediction models. To deliver manually gathered data we have a maximum of three individuals. Additionally, registering every data source manually is impossible for data sources like the amount of Bluetooth signals in the surroundings.

An alternative way of scoring is to choose for the determination of a relative score for each data point using the rest of the test group as a reference point. A disadvantage is that the scores cannot be related to real data and are therefore useless as comparison material for future studies about general smartphone behavior. Within this study however, the use of relative scores for each data point gives the advantage that comparison between one data point and another can be done fast and efficient if the same scale is maintained (e.g. a score between 0 and 100), while scores can be used as complements or as counterweights when creating a social profile. For instance, this would enable sociability to be captured as one factor, which would speed up decision-making within the healthcare sector.

To finish the scoring process, the experts suggested adding some basic statistics describing the total population to create some contextual information for the individual scores. As basic statistics we add the following: the population average, standard deviation, confidence interval, range and the total sample size.

5. Results

The BeHapp application registered a total of 13.688 events cumulatively for 10 individuals over an average of 11.35 days, which together formed the final dataset on which further analysis is performed. To discover any statistical relationships between social acts, linear regression analysis was performed for both the real values and the associated scores. The data points that were analyzed included all social acts in every of the three different dimensions (if relevant). Only one weak relationship has been found between CO_frequency and CM_frequency, which after filtering of extreme values remained a significant relationship ($p = 0.024^{*}$; $r^2 = 0.539$; $\beta = 0.734$). This means that the average occurrence of outgoing (CO) and missed phone calls (CM) per day is correlated. To discover any significant relationships between scores, several linear regression analyses were performed creating different matrices. The strongest relationship was found between the average incoming and outgoing communication score ($p = 0.015^{*}$; $r^2 = 0.541$; $\beta = 0.735$). The remaining matrices showed some relationships for CI/CO, CI/CI_duration, CO/CI_duration and CO_duration/CO_diversity, but in all cases, the relationships are strongly influenced by extreme values, which after removal leave uncorrelated collections of data points.

5.1. Confounding factors: personality and smartphone use

In order to determine the influence of a user's personality on his smartphone behavior, regression analysis is performed for each of the Big Five personality traits, the social acts and all available scores (both separated and aggregated). Concerning social acts, only one correlation has been found between inquisitiveness and the diversity in outgoing SMS message receivers, which is considered a strong relationship ($p = 0.030^{**}$; $r^2 = 0.834$; $\beta = 0.913$). When linked to any of the calculated scores, two correlations have been found for personality traits: the score for the amount of outgoing SMS messages is positively correlated with extraversion ($p = 0.097^{*}$; $r^2 = 0.307$; $\beta = 0.554$), and a higher outgoing communication score is associated with higher inquisitiveness ($p = 0.087^{*}$; $r^2 = 0.322$; $\beta = 0.568$). However, in both cases the model can only explain about 30% percent of the variance, which decreases the potential of the prediction model. Finally, two significant relationships have been found including personality traits only. First, we found extroversion and accommodation to be strongly related ($p = 0.031^{**}$; $r^2 = 0.459$; $\beta = -0.678$). Secondly, orderliness and inquisitiveness show a weak relationship ($p = 0.086^{*}$; $r^2 = 0.324$; $\beta = -0.569$), but again the value of r-squared indicates a weak fit of the prediction model in both cases. The findings of this model should be interpreted with caution, since they are based on five subjects, which limiting the overall reliability of predictions.

To discover the role of smartphone use in the expression of an individual's sociability, we examine the statistical relationships between the created smartphone use score and the other scores, both separated and aggregated. The smartphone use score (which is equal to application activity) shows a negative statistical relationship with the incoming ($p = 0.021^{**}$; $r^2 = 0.505$; $\beta = -0.711$), outgoing ($p = 0.028^{**}$; $r^2 = 0.475$; $\beta = -0.689$) and total communication score ($p = 0.012^{**}$; $r^2 = 0.566$; $\beta = -0.752$). Maintaining the significance levels $p < 0.1^{*}$, $p < 0.05^{**}$ and $p < 0.01^{***}$, all these relationships can be considered strong. However, after the removal of the extreme values for person 30, all previous found correlations are gone. Therefore, we assume for this research the effect of smartphone use on overall sociability to be too small to be of influence.

5.2. Sociability profiling

The final sociability score for each test subject is visualized in Fig. 2. As shown by the lines in the graph the average social exploration score is found to be 47.87 and the average communication score is 49.75. Note that person 23 is omitted while we were unable to create a social exploration score due to insufficient GPS and Bluetooth signal data. As can be seen in Fig. 2, the differences between the subjects in sociability scores are marginal as the subjects all cluster around the center. Still a distinction can be made between subjects that score either below or above the averages. Four subjects score below average on both social exploration and communication score (person 18, 19, 21 and 30), one subject scores only below the social exploration score average (person 20) and two subjects score only below the average on communication (person 16 and 22). Two subjects score above the average for both social exploration and communication which are person 25 and 26.

As a final result, Table 1 presents a summary of the highest-level score that will function as the final sociability profile of a subject, which is in this case for person 16. An overview of all the underlining scores and all sociability scores for the other test subjects are added to appendix A.

6. Conclusion

For the creation of a social profile, the first step encloses the determination of the social profile contents. Maintaining the term sociability and the corresponding definition by Cheek & Buss (1981), we came to the distinction of the first two categories: communication and physical social exploration. During the business understanding phase, the expert interviews in accordance with the results from the systematic literature review led to the definition of the following category dimensions; for communication, the dimensions frequency, diversity and duration were defined and for social exploration, the dimensions visit duration, visit distance, social density and movement pattern diversification. These categories and dimensions later formed the basis for the sociability model. The next step was to examine the smartphone data in the data understanding phase and select, transform and cleanse the data to create a final data set, which included phone call data, application activity, GPS events and Bluetooth signal events. Then, the data sources needed to be linked to the previously defined categories and dimensions in correspondence with the concept of sociability: a result which can be found in the completion of the sociability model. Subsequently, a scoring mechanism was constructed to assign values to the different dimensions of sociability and to come to an eventual sociability score by aggregation of underlying scores. In addition to the sociability score, the influence of possible confounding factors was evaluated, and if relevant, added to the list of final scores to form the eventual sociability profile. For this matter, neither personality nor smartphone use appeared to influence sociability enough to conclude that either should be included in the final sociability profile. As a test case, we subsequently applied the sociability model on data collected from a group of 10 students to assess the usefulness of the eventual model. To limit the influence of other confounding factors, we picked test subjects based on the same user characteristics (age, gender, education level, mental condition and level of smartphone experience). The results from this test case reveal a weak relationship between outgoing calls and missed calls, which may be caused by the fact that people who miss a phone call are likely to return the call to discover the reason of the first call. Furthermore, one weak relationship has been found between the incoming communication score and outgoing communication score, which could indicate that a person's communication profile can be reflected by both the incoming and outgoing communication. Finally, the sociability profiles of the 10 students were successfully created using the sociability model, which shows that the model is a possible answer for the main question. This pilot study provided insight in the usability of the individual sociability scores for future smartphone application to provide longitudinal objective measures of normal and atypical human social behavioral profiles in their natural environment.

7. Discussion, limitations and future research

For this research we distinguish two types of limitations; instrumental limitations, which are concerned with the restrictions imposed by the BeHapp application and general research design limitations. We will expand both types in the following section below.

The instrumental limitations are the consequence of the BeHapp application still being in the alpha development phase. Several test reports were written during this research with additional requirements, but the version update arrived too late to use the latest

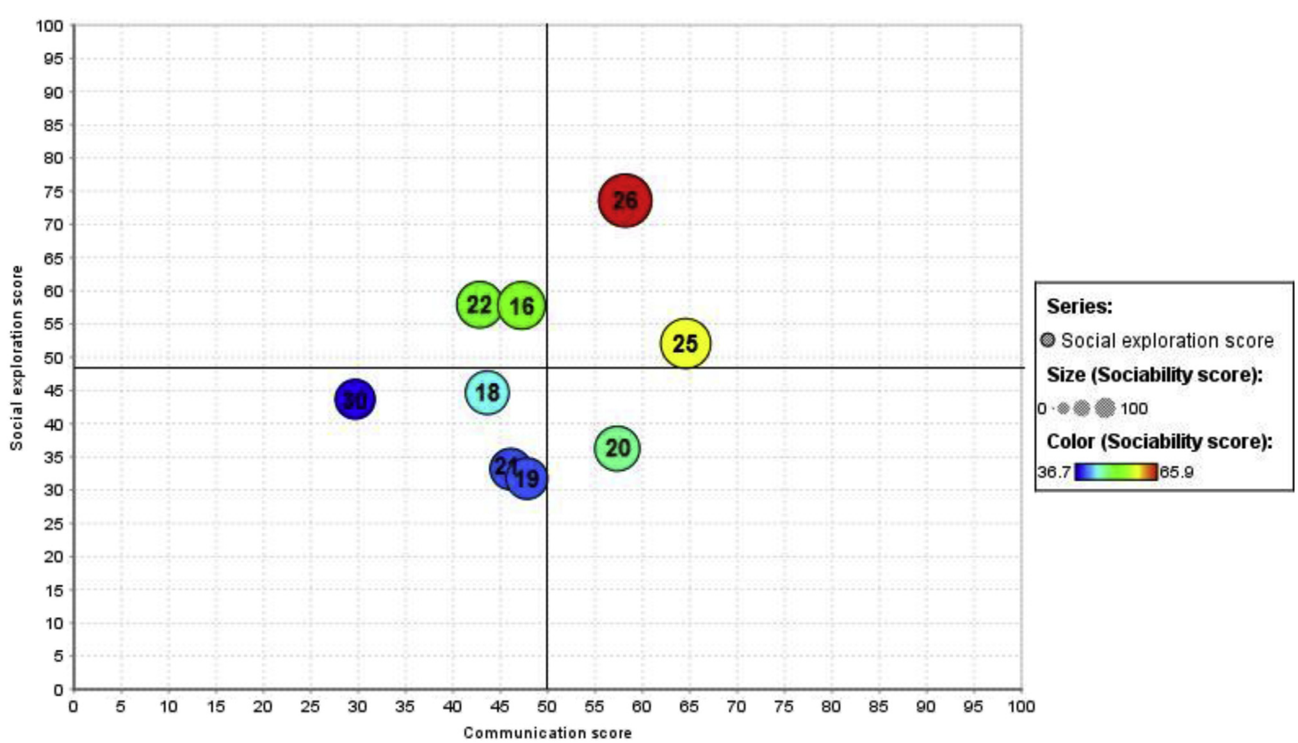

Fig. 2. Sociability score per subject.

Table 1
The Sociability profile of test subject 16.

Test subject 16	Score	Average	Standard deviation	Sample size	Lower bound	Upper bound	Min	Max	Range
Communication score	47.23	49.75	9.85	10	43.65	55.85	29.67	64.56	34.89
Social exploration score	57.80	47.87	12.98	9	39.39	56.35	31.73	73.58	41.85
Sociability score	52.51	49.43	9.29	10	43.67	55.19	36.67	65.88	29.20

version from the start. The first instrumental limitation encloses the deficient data retrieval from several data sources including for instance WhatsApp messenger, Facebook and Twitter. Therefore, these data sources could not be tested and for now cannot play a role in the creation of a sociability profile. Also, bugs in the alpha version caused the application or even the entire smartphone to crash at certain times, which forced several test subjects to drop-out if the smartphone became inoperable. In replacement, new test subjects had to be found to fill the empty spaces which caused research delay and led to differential observation periods of test subjects. Another disadvantage of the first version is the rapid battery depletion which discouraged some test subjects to use their smartphone.

The first research design limitation is the sample size ($n = 10$), which can be explained by the fact that it is difficult to find individuals that are willing to give up some of their privacy. This resulted in the specification of test subject characteristics to ensure internal validity. A consequence is that only a small part of the population could be represented, leaving room for additional research to examine for instance differences across gender and education level. The second limitation is the time restriction; participants where observed over a time period of two weeks, which raises the question whether the same results hold over a longer period of time. Furthermore, there is the diversity in smartphone type the participants own, varying from Samsung, Sony, HTC and LG devices. As performance testing is not performed strictly for each of these devices, it is unknown whether differences exist in the registration of events. As an example, it could be the case that for some devices, the Bluetooth component can register more Bluetooth-enabled devices in one scan than the Bluetooth component of other devices.

7.1. Future research

This study is a pilot study for the actual study which is aimed for to test the method in a subsample of an ongoing longitudinal youth cohort, involving patients diagnosed with a form of schizophrenia, from which conclusions will be derived about the usefulness of the current model version in a practical, clinical matter. Finally and ideally, the validated method can formally be deployed in a clinical context for profiling patients and subsequently using this social exploration profile as an addition for the diagnoses of several mental illnesses. However, the results from this research can also be extended in several other directions. We will walk through them based on the size of impact in ascending order.

First, the sociability model can be complemented with new variables from new data sources. Acts on social media for instance can be added, as presumptions arise from existing literature mentioning that relationships exist between extraversion, social anxiousness, loneliness and Facebook use (Aspendorpf & Wilpers, 1998; Ebeling-Witte, Frank, & Lester, 2007; Ryan & Xenos, 2011). Other additional data sources include text-messenger applications like WhatsApp (500 + m users worldwide), WeChat (438 + m users worldwide) and Line (400 + m users worldwide) (TNW, 2014; Forbes, 2014), which form an essential part of modern communication and can be added as new communication variables both in the frequency and diversity dimension.

Another way of adding new variables is exploiting data sources in innovative manners. GPS data for instance can be exploited in several different ways, creating scores based on location visits, movement patterns or travel analysis. For movement patterns, a presumption about variation was mentioned during the expert interviews in which one of the experts states that he expects more variety in movement patterns among schizophrenic patients, as they have the tendency to avoid seemingly crowded places. One challenge is to select the most appropriate techniques given the configuration of variable types, which we aim to explore from a meta-algorithmic perspective (Vleugel, Spruit, Van Daal, 2010; Spruit, Vroon, & Batenburg, 2014).

When adding new variables, it could also be the case that some variables need to be replaced, which could be the case with social density. Other data sources are available like WiFi-signal density and GPS-data linked to a region density database, which could be complementary or replacing factors for the concept of social density.

A final interesting concept mentioned in the literature is the use of social rhythms. Using multiple data sources, such as the location, the identification of Bluetooth signals and the time, rhythms can be discovered that describe patterns in a user's life. When focusing on anomalies in these social rhythms, social withdrawal or perhaps even mental illness exacerbation can be identified more easily when observing scores over an extended period of time.

Another direction for future research would be the improvement of the model. A good start would be the addition of weights to the existing and/or newly added factors. For this research, the aggregated scores are created by taking the average of other scores, maintaining the assumption that each score has an equality in weight, while for now we have no reason to presume that some variables are more important than others within this aggregation (e.g. the frequency of conversations may be of more significance than the duration of conversations). In a follow-up study these weights can be established, when comparing the scores of this study's test group with the scores of a group of people diagnosed with social deficits. This will show that some differences in particular scores between the groups may be more significant than others or that some scores may even be useless. Subsequently, the scores can be weighted based on the significance of the difference to create the optimal balance when creating an aggregated score.

Although no evidence was found to prove a relationship exists between personality and sociability, our preliminary data suggest that personality profiles can be observed, which is of interest because several other studies did find relationships between personality and smartphone usage. In case these relationships exist, one could include these correlations into the analysis by using the strength of the relationship as an explanatory factor for the sociability score. Personality then can be integrated in the model by reducing the sociability score equal to the influence of the personality traits on the separated communication or social exploration scores.

The third and final suggestion for future research is scaling the usefulness of the model by creating a baseline for a more diverse test group, including for instance differences across gender, age, culture, educational level and profession. In this way, potential patients can be compared with a more diverse group making eventual diagnoses more reliable.

Appendices

Appendix A: Separated scores

	16	18	19	20	21	22	23	25	26	30
Frequency										
Incoming call	34,39	38,68	48,37	47,49	52,36	50,69	84,90	75,44	35,32	32,38
Outgoing call	35,97	52,67	57,78	42,23	46,23	49,21	44,29	95,82	35,36	40,44
Incoming sms	34,68	50,27	70,94	42,84	85,61	45,20	62,08	40,53	31,73	36,12
Outgoing sms	37,38	37,38	37,38	37,38	88,34	49,51	63,88	66,95	37,38	44,41
Application Activity	64,49	47,28	34,00	35,59	39,26	70,63	47,29	35,16	40,99	85,31
Missed call	28,08	65,17	54,98	40,79	33,68	68,51	61,21	76,36	28,08	43,15
Diversity										
Incoming call (div)	67,57	50,36	67,57	67,57	44,62	35,29	41,75	41,75	67,57	15,93
Outgoing call (div)	51,00	40,15	42,86	58,24	44,49	36,53	57,52	38,26	94,42	36,53
Incoming SMS (div)	71,50	37,75	36,07	71,50	36,22	52,52	50,79	51,25	20,88	71,50
Outgoing SMS (div)	34,10	34,10	34,10	34,10	58,58	76,53	57,97	70,47	34,10	65,93
Missed call (div)	22,87	54,33	44,89	66,91	66,91	48,03	66,91	39,38	22,87	66,91
Duration										
Incoming call (dur)	x	38,29	38,50	44,08	39,47	37,98	70,85	80,83	x	x
Outgoing call (dur)	x	41,51	31,71	84,31	49,33	47,29	62,01	55,24	x	28,58
Social exploration										
BT	71,16	31,61	x	x	22,93	53,72	x	65,62	61,30	43,67
Distance	44,43	57,74	31,73	36,24	43,42	62,08	x	38,50	85,87	x

Aggregated scores

	16	18	19	20	21	22	23	25	26	30
Communication score										
Communication_in	50,98	42,44	51,48	53,05	45,48	41,32	65,83	66,00	51,45	24,15
Communication_out	43,49	44,78	44,12	61,59	46,69	44,34	54,61	63,11	64,89	35,19
Communication score	47,23	43,61	47,80	57,32	46,08	42,83	60,22	64,56	58,17	29,67
Social exploration score										
SE score	57,80	44,67	31,73	36,24	33,17	57,90	x	52,06	73,58	43,67

Sociability scores

Sociability	16	18	19	20	21	22	23	25	26	30
Sociability score	52,51	44,14	39,76	46,78	39,63	50,37	60,22	58,31	65,88	36,67

References

Adams, B., Phung, D., & Venkatesh, S. (2008). Sensing and using social context. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), 5*(2), 11.

Aspendorpf, & Wilpers. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology, 74*(6), 1531–1544.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, N.J: Prentice-Hall.

Barney, M., & McCarty, T. (2003). *The new six sigma: A leader's guide to achieving rapid business improvement and sustainable results*. Upper Saddle River, N.J: Prentice Hall PTR.

Bierlaire, M., Chen, J., & Newman, J. (2010). *Modeling route choice behavior from smartphone GPS data*. Report TRANSP-OR, 101016, 2010.

Buijs, M., & Spruit, M. (2015). Determining the relative importance of webpages based on social signals using the social score and the potential role of the social score in an asynchronous social search engine. In A. Fred, J. Dietz, D. Aveiro, K. Liu, & J. Filipe (Eds.), *Knowledge discovery, knowledge engineering and knowledge management - 6th International joint conference, IC3K 2014, Rome, Italy, October 21-24, 2014, Revised selected papers* (pp. 118–131). ScitePress.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., et al. (2000). *CRISP-DM 1.0 step-by-step data mining guide*.

Cheek, J. M., & Buss, A. H. (1981). Shyness and sociability. *Journal of Personality and Social Psychology, 41*(2), 330–339.

De Reuver, M., Bowman, H., Heerschap, N., & Verkasalo, H. (2012). Smartphone measurement: do people use mobile applications as they say they do?. In *International conference on mobile business. Paper 2*.

Ebeling-Witte, Frank, & Lester. (2007). Shyness. Internet use, and personality. *Cyberpsychology & Behavior, 10*(5), 713–716.

Ernst, & Young. (2011). *Onderzoeksresultaten ICT barometer over ICT en gezondheidszorg*. Retrieved from http://www.ey.com/Publication/vwLUAssets/ICT_Barometer_-_Nederlander_wil_beter_gebruik_van_ICT_in_gezondheidszorg/$FILE/ICT%20Barometer%20over%20ICT%20en%20gezondheidszorg%207%20december%202011.pdf.

Forbes. (2014). *Whatsapp hits 500 million users*. Retrieved October 19, 2014, from http://www.forbes.com/sites/amitchowdhry/2014/04/22/whatsapp-hits-500-million-users/.

Grace, J. H., Zhao, D., & Boyd, D. (2010). Microblogging: what and how can we learn from it?. In *CHI 2010 Extended Abstracts on Human Factors on Computer Systems (CHI EA 2010)* (pp. 4517–4520). Atlanta, Georgia, USA.

Hightower, J., & Boriello, G. (2001). A survey and taxonomy of location systems for ubiquitous computing. *Computer, 34*(8), 57–66.

Kaikkonen, A. (2008). Full or tailored mobile web- where and how do people browse on their mobiles?. In *International Conference on Mobile Technology, Applications, and Systems (Mobility '08), Article 28*. New York: ACM Press.

Maheshwaree, P., Ylä-Jääski, A., Verkasalo, H., & Hämmäinen, H. (2009). *Determinants of smartphone usage* (Unpublished manuscript).

Meulendijk, M., Meulendijks, E., Jansen, P., Numans, M., & Spruit, M. (2014). What concerns users of medical apps? Exploring non-functional requirements of medical mobile applications. In *22nd European conference on information systems (ECIS). Tel Aviv, Israel*.

Nicolai, T., & Kenn, H. (2006). Towards detecting social situations with Bluetooth. In *Adjunct proceedings Ubicomp*.

Pachidi, S., Spruit, M., & Van der Weerd, I. (2014). Understanding users' behavior with software operation data mining. *Computers in Human Behavior, 30*, 583–594.

Ryan, T., & Xenos, S. (2011). Who uses Facebook? an investigation into the relationship between the big Five, shyness, narcissism, loneliness, and facebook usage. *Computers in Human Behavior, 27*(5), 1658–1664.

Shearer, C. (2000). The CRISP-DM Model: the new blueprint for data mining. *Journal of Data Warehousing, 5*(4).

Spruit, M., Vroon, R., & Batenburg, R. (2014). Towards healthcare business intelligence in long-term care: an explorative case study in the Netherlands. *Computers in Human Behavior, 30*, 698–707.

Steg, L., Buunk, A. P., & Rothengatter, T. (2008). *Applied social Psychology: Understanding and managing social problems*. Cambridge, UK: Cambridge University Press.

Straka, R. J., Fish, J. T., Benson, S. R., & Suh, J. T. (1997). Patient self-reporting of compliance does not correspond with electronic monitoring: an evaluation using isosorbide dinitrate as a model drug. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 17*(1), 126–132.

TNW. (2014). *Chat app line passes 400 million registered users, inks record 10 billion messages sent in one day*. Retrieved October 19, 2014, from http://thenextweb.com/apps/2014/04/02/chat-app-line-passes-400-million-registered-users-inks-record-10-billion-messages-sent-in-one-day/.

Vleugel, A., Spruit, M., & Daal van, A. (2010). Historical data analysis through data mining from an outsourcing perspective: the three-phases method. *International Journal of Business Intelligence Research, 1*(3), 42–65.

Yan, Z., Yang, J., & Tapia, E. M. (2013, September). Smartphone bluetooth based social sensing. In *Proceedings of the 2013 ACM conference on pervasive and ubiquitous computing adjunct publication* (pp. 95–98). ACM.