CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Factitious disorder and its online variant Munchausen by Internet: understanding motivation and its impact on online users to develop a detection method
Authors	Lawlor, Aideen
Publication date	2018
Original Citation	Lawlor, A. 2018. Factitious disorder and its online variant Munchausen by Internet: understanding motivation and its impact on online users to develop a detection method. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2018, Aideen Lawlor. - http://creativecommons.org/licenses/by-nc-nd/3.0/
Download date	2026-11-12 13:16:10
Item downloaded from	https://hdl.handle.net/10468/7744

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Factitious Disorder and its online variant Munchausen by Internet: Understanding motivation and its impact on online users to develop a detection method.

By

Aideen Lawlor

A thesis by publication submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy (PhD) to the National University of Ireland, Cork.

2018

School of Applied Psychology

Supervisor: Dr Jurek Kirakowski

Head of School: Dr John McCarthy

CONTENTS

Statement of original authorship i
Acknowledgements ii
Dedication iii
Abstract iv
List of publications by candidate vi
CHAPTER 1.0 INTRODUCTION 1
1.1 Impetus for the research 2
1.2 Factitious Disorder 6
1.2.1 Diagnosis 6
1.2.2 Symptoms 8
1.2.3 Prevalence 8
1.2.4 Characteristics 9
1.2.5 Motivation 9
1.2.6 Treatment 10
1.2.7 Factitious Disorder imposed on another (Munchausen by proxy) 11
1.2.7.1 Diagnosis 11
1.2.7.2 Presentation and prevalence 12
1.2.7.3 Characteristics 13
1.2.7.4 Motivation 13
1.2.7.5 Treatment 14
1.2.8 Impact on those duped by FD and FDIA victims 14
1.2.9 Stigma surrounding FD 16
1.3 Munchausen by Internet 17
1.3.1 Case study of MbI 17
1.3.2 Overview of MbI 20
1.3.3 Motivation 21
1.3.3.1 MbI 21
1.3.3.2 MBPBI 22
1.3.3.3 MBIBT 24
1.3.3.4 Malingering 24
1.3.4 Issues with applying extrapolations from FD to MbI 25
1.3.5 Impact of MbI on those duped 25
1.3.6 Detecting and managing MbI 27
1.4 Digital deception 28

1.4.1 Advantages of the online environment 28
1.4.2 Types of online identity deception 30
1.4.3 Deceptive tactics 31
1.4.4 Identifying online identity deception 32
1.5 Social Language Processing (SLP) 33
1.5.1 Text classification applied to online deception. 35
1.5.1.1 Opinion spam 35
1.5.1.2 Online harassment 36
1.5.1.3 Sockpuppets 37
1.5.1.4 Sexual predators 39
1.6 Research Aims 40
1.7 Qualitative Research Methodology 43
1.7.1 Grounded Theory Overview 44
1.7.2 Basic tenets of Grounded Theory 44
1.7.3 Glaserian vs. Straussian 46
1.7.4 Objectivist vs. Constructivist 47
1.7.5 Suitability of Grounded Theory 50
1.7.6 Three coding phases of Constructivist Grounded Theory 53
1.7.6.1 Open Coding – Identifying Codes and concepts 54
1.7.6.2 Focused Coding – Identifying categories 55
1.7.6.3 Theoretical Coding – Identification and integration of core categories 55
1.7.6.4 Quality Assurance 56
CHAPTER 2.0 When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder 59
2.1 Abstract 60
2.2 Introduction 61
2.2.1 Revision of DSM criteria 61
2.2.2 Dependence on case studies 63
2.2.3 Motivation and aetiology 63
2.2.4 Symptoms 64
2.2.5 Aims of the study 64
2.3 Method 65
2.3.1 Sample 65
2.3.2 Limitations of sample 65
2.3.3 Data collection 66
2.3.4 Ethical consideration 66

2.3.5 Data analysis 67
2.4 Results 73
2.4.1 The characteristics of FD among members 74
2.4.1.1 Early onset 74
2.4.1.2 Feigned and genuine disorders 74
2.4.1.3 FD behaviour 75
2.4.2 Motivation 77
2.4.2.1 Long-term motivation 77
2.4.2.2 Affection 77
2.4.2.3 Difficult childhood 78
2.4.2.4 Enjoyment 79
2.4.2.5 Coping mechanism 80
2.4.2.6 Sense of achievement (duping delight) 80
2.4.2.7 Validation of genuine disorder 81
2.4.2.8 Identity issues 81
2.4.2.9 External gains 81
2.4.2.10 No reason for behaviour 81
2.4.3 Episodic triggers 82
2.4.4 Symptoms 83
2.4.4.1 Poor mental health 84
2.4.4.2 Dissociation from real self 85
2.4.4.3 Guilt and shame 85
2.4.4.4 Reduction in quality of life 85
2.4.4.5 Negative self-perception 86
2.4.4.6 Addiction 86
2.4.4.7 Isolation 87
2.4.4.8 Ostracised 87
2.4.5 Treatment 87
2.4.5.1 Barriers to treatment 88
2.5 Discussion 93
2.5.1 Motivation and aetiology 94
2.5.1.1 Substantiated motivational theories and aetiology 94
2.5.1.2 Is motivation conscious or unconscious? 95
2.5.1.3 Cursory comparison between behavioural addiction and FD 96
2.5.2 Symptoms 97
2.5.3 Barriers to seeking treatment 98

2.5.4 Recovery strategies 99
2.6 Limitations 100
2.7 Conclusion 100
CHAPTER 3.0 Claiming someone else's pain: A Grounded theory analysis of online community user's experiences of Munchausen by Internet. 102
3.1 Abstract 103
3.2 Introduction 104
3.2.1 Online communities and trust 104
3.2.2 Online Communities and Identity Deception 105
3.2.3 Munchausen by Internet 105
3.3 Method 109
3.3.1 Data collection 109
3.3.2 Ethical consideration 110
3.3.3 Data analysis 111
3.4 Results 119
3.4.1 Perceived motivation 119
3.4.1.1 Attention and Sympathy 119
3.4.1.2 Maliciousness 120
3.4.1.3 Mental Illness 120
3.4.1.4 Material exploitation 121
3.4.2 Personal effect of MbI 121
3.4.2.1 Emotionally exploited and manipulated 121
3.4.2.2 Unperturbed 122
3.4.3 Effect of MbI on community 122
3.4.3.1 Community as a support group 122
3.4.3.2 Air of suspicion 122
3.4.3.3 Witch hunts 123
3.4.3.4 Resilience 124
3.4.4 Detecting MbI 124
3.4.4.1 Types of MbI 124
3.4.4.2 Clues which raise suspicion 125
3.4.4.3 Clues associated with genuine users 127
3.4.4.4 Clues which confirm suspicions 127
3.4.5 Managing MbI 128
3.4.5.1 Difficult for moderators to control 128
3.4.5.2 Publicly outing MbI users 129

3.4.5.3 Report and ignore 130
3.4.5.4 Prevention through caution and awareness 130
3.5 Discussion 131
3.5.1 Portrayal of the 'ideal victim' 131
3.5.2 Unexpected sympathy towards MbI users 132
3.5.3 Self-censorship limiting opportunities to confer communal normality 133
3.5.4 Difficulty identifying MbI users 134
3.5.5 No deterrents for MbI 135
3.5.6. Preventative action 136
3.6 Conclusion 137
CHAPTER 4.0 Detecting Munchausen by Internet: Development of a Text Classifier through Machine Learning. 139
4.1 Abstract 140
4.2 Introduction 141
4.2.1 Social Language Processing (SLP) paradigm 142
4.2.2 Stage 1: Identifying potential language features 142
4.2.2.1 deception theories 142
4.2.2.3 Using LIWC to create a linguistic profile of deception 148
4.2.2.4 Munchausen's profile 155
4.3 Method 157
4.3.1 Creating the data set 157
4.3.2 Stage 2: Automatically extracting the relevant discourse features 159
4.3.2.1 Pre-processing the data set 159
4.3.2.2 Feature extraction 160
4.3.3 Stage 3: Developing statistical classifiers 161
4.3.3.1 Classifier Training 161
4.4 Results 164
4.4.2 Results from LIWC2015 model 164
4.4.3 Evaluation of N-gram model 167
4.5 Discussion 170
4.5.1 Predictive LIWC dimensions 171
4.5.1.1 Captivating readers for their attention and sympathy 171
4.5.1.2 Social outlet 172
4.5.1.3 Relation to LBC theories of deception 173
4.5.2 Predictive words in the N-gram model 174
4.6 Limitations 175

4.7 Conclusion 176
CHAPTER 5.0 Conclusion 178
5.1 Introduction 178
5.2 Practical application value 179
5.2.1 Limitations and future directions 180
5.3 Theoretical implications 181
5.3.1 New theories of FD 181
5.3.1.1 Limitations and directions for future research 184
5.3.2. Dynamics of online communities 185
5.3.2.1 Limitations and directions for future research 188
5.3.3 Using the SLP paradigm to identify LBC's of deception 189
5.3.3.1 Limitations and directions for future research 190
5.4 Conclusion 191
CHAPTER 6.0 Consolidated list of references 193

STATEMENT OF ORIGINAL AUTHORSHIP

I hereby certify that the submitted work is my own work, was completed while registered as a candidate for the degree stated on the Title Page, and I have not obtained a degree elsewhere on the basis of the research presented in this submitted work.

Signed: _____ Date: _____

(Aideen Lawlor)

`

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis supervisor Dr Jurek Kirakowski for his continuous support throughout and the wealth of knowledge he contributed. His patience and motivation gave me the confidence to freely explore new methodologies and lines of research, none of which would have been achievable without his support.

I would also like to thank the School of Applied Psychology for the studentship I received, not only did it provide financial support but was also an invaluable opportunity to learn more about the world of academia.

I would also like to thank my family and friends: my parents, my brothers and Wiola Stępniak for their emotional support throughout my thesis research and my life in general.

DEDICATION

I would like to dedicate this thesis to my mother, Yvonne, and my father, James. For their unwavering and unconditional love and support throughout my life, I am eternally grateful.

ABSTRACT

The overarching aim of the research in this thesis was to develop a method of detecting Munchausen by Internet (MbI) and garner an understanding of the dynamics of online communities faced with MbI. Ground work studies were required to learn more about the disorder, to decide exactly what method of detection would be most appropriate. This involved a review of the existing literature available on MbI *(paper 1; Munchausen by Internet).* It also involved conducting two studies which focused on experiences from the perspective of those with Factitious Disorder (FD) *(paper 2; When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder)* and MbI from the perspective of victims *(paper 3; Claiming someone else's pain: A Grounded theory analysis of online community user's experiences of Munchausen by Internet).* Both these studies were necessary as FD and its online variant MbI are some of the most poorly understood and under researched pathologies. This is primarily because of the difficulty in obtaining and retaining participants who have experience of the disorder. Therefore, what was previously known about the disorder was largely speculative. The research presented in this thesis overcame the issue of recruitment and retentions of participants, by analysing the first-hand accounts written online by those who have experience of the disorders. The information obtained from the two groundwork studies was used in the third study to decide on and develop an appropriate method of detecting MbI and for interpreting the discriminate attributes *(paper 4; Detecting Munchausen by Internet: Development of a Text Classifier through Machine Learning).*

Beyond applying the findings of these studies to the development of the classifier, they also made new theoretical contributions to the existing literature on FD and MbI. The first two studies provide the very first large-scale studies of FD and MbI, using first-hand accounts from those it directly affects rather than observations that are speculative. Grounded Theory was used to analyse the text as it does not require an a priori theoretical framework but allows the data to build the theoretical framework itself, resulting in more innovative findings. The findings offer a new perspective of FD, one which contrasts with traditional theories and indicates that FD may be closely aligned with addiction. The second study examined the dynamics within an online community faced with MbI. The primary findings were that MbI users were targeting

‘ideal victim’ persona which offered protection from suspicion and increased the level of attention and sympathy they could receive. The presence or possible presence of MbI also resulted in members of online communities using strategies to avoid false accusations or being duped. These strategies had the unfortunate consequence of potentially eroding the therapeutic benefits of online communities, in particular personal empowerment, by restricting opportunities to confer normality and cultivate interpersonal support. In addition, the methods used by online community members and their moderators to detect MbI were uncovered. It typically involved high-level deception cues which raised suspicions and the checking authoritative references to confirm or refute these suspicions.

The findings from study one and two, as well as the literature review from paper one, offered no overt cues which could be consistently attributed to MbI and offered no support for the feasibility of psychometric testing to detect MBI. Therefore, it was decided that covert deception required a covert method of detection. To this end the SLP (Social Language Processing) framework, which integrates psychology and computer science, was applied to develop a text classifier through machine learning algorithms. This covert method has already been successfully used to detect written deception online. Two text classifiers were developed in study three using Linguistic Inquiry Word Count (LIWC2105) dimensions and n-grams obtained from a bag-of words model, with respective prediction accuracies of 81.11% and 81.67%. These classifiers added a practical application value to the research conducted in this thesis, by producing a method of detecting MbI that can be used by moderators and as a vetting and investigative tool for internet mediated researchers. There were also theoretical contributions obtained from study three. Some of the discriminate attributes used by the classifiers appeared to be unique to Munchausen’s and were associated with the motivation for the behaviour, which supports the growing move towards domain specificity when interpreting Linguistic Based Cues (LBC) of deception. The remaining LBC’s of deception concurred with established deception theory, particularly reduction of cognitive complexity. Overall the research described in this thesis has made new contributions to the existing theories surrounding Factitious Disorder (FD), MbI and Linguistic Based Cues (LBC’s) of deception. It also has a practical application value by creating a classifier which differentiates between text written by genuine people and those exhibiting Munchausen’s.

LIST OF PUBLICATIONS BY CANDIDATE

Published journal papers

1. Lawlor, A., & Kirakowski, J. (2014). When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder. *Psychiatry research, 218*(1-2), 209-218.
2. Lawlor, A., & Kirakowski, J. (2017). Claiming someone else's pain: A Grounded theory analysis of online community user's experiences of Munchausen by Internet. *Computers in Human Behaviour, 74,* 101-111.

Journal manuscripts under review or submitted for review

1. Lawlor, A. & Kirakowski, J. (2016). Detecting Munchausen by Internet: Development of a Text Classifier through Machine Learning. *Manuscript submitted for review in Computers in Human Behaviour.*

Book Chapters under review

1. Lawlor, A. (2017). Munchausen by Internet. *Manuscript with publishers for the Wiley Handbook of Human-Computer interaction (Vol 2).*

CHAPTER 1.0 INTRODUCTION

Aideen Lawlor

School of Applied Psychology,
University College Cork

Chapter submitted for review:
Wiley Handbook of Human-Computer interaction (Vol 2)

Statement of contributions

Lawlor, A. (Candidate)
Writing and compilation of manuscript.

Section 1.3. MbI is an altered version of the paper referred to above. For the purposes of the literature review it excludes references to the work conducted by Lawlor and Kirakowski (2014), (2016a), (2016b) which is presented in the remaining chapters.

1.1 IMPETUS FOR THE RESEARCH

McLaughlin et al. (2016) describes Online Support Groups (OSG's) as nonprofessional groups of peers who meet on various platforms (blogs/social media/discussion forums/live chats etc.) available online, who are unified by their shared experiences, identities or illness and are not constrained by space (meeting face-to-face) or time (asynchronous exchanges). OSG's have been championed as a valuable resource for those who are coping with difficult life events as they offer the same types of supports that are available offline including informational, esteem, emotional, tangible and network (Braithwaite et al., 1999; Preece, 1999; McCormack & Coulson, 2005; Coulson et al., 2007; Coursaris & Liu, 2009; Coulson & Greenwood, 2012). The most frequent type of support experienced in OSG's are informational and emotional while the least are tangible and network (Braithwaite et al., 1999). Furthermore, OSG's may be particularly appealing to those who are experiencing self-stigma associated with their physical or mental health issues and are having difficulty discussing these issues offline (Lawlor & Kirakowski, 2014; Mo & Coulson, 2010 Coursaris & Liu, 2009). This is because OSG's do not require a person to reveal their real identity and the shroud of anonymity afforded online encourages increased self-disclosure, because they are not exposed to the same risks that would accompany self-disclosure offline (Joinson, 2001; McKenna, 2008). OSG's also offer supplementary support to those who are seeking support offline, are geographically isolated from their peers and/or require 24/7 instant access to support e.g. those recovering from addiction (Stommel, 2009; Fenner & Gifford, 2014). Physical and mental health organisations are also including OSG spaces within their websites to facilitate peer support (Ware, 2011).

The accessibility and easy availability of support has likely led to the growing popularity of OSG's (Nimrod, 2012a). Approximately 7% of internet users participate in OSGs (Fox & Purcell, 2010). However, the question remains as to whether OSG's produce any beneficial outcomes for those who participate in them. Two systematic reviews have been conducted to assess the efficacy of OSG's. The first was by Eysenbach, Powell, Englesakis, Rizo and Stern (2004) who examined 38 studies that evaluated the health or social benefits of OSG participation. These studies employed a variety of methods including control groups, before and after design, interrupted

time series and cohort studies: and the majority used depression and social support as their outcome measures. Eysenbach et al. (2004) concluded that there was no evidence to show that OSG's were either beneficial or harmful to users. This finding was partially attributed to the litany of issues encountered with trying to obtain studies that were evaluating OSG's that were purely peer to peer. Instead most involved health professionals or were used as a supplement to more complex interventions. Furthermore, the studies were of poor quality e.g. improper randomisation methods and allocation concealment. Griffiths et al. (2009) conducted a second systematic review which focused on depressive symptoms as the outcome measure. However, the same issue encountered by Eysenbach et al. (2004) were also reported by Griffiths et al. (2009), poor design quality of the studies. The importance of this issue was reinforced further as there was a strong association between the positive outcomes of a study and low design quality. Griffiths et al. (2009) concluded that the lack of evidence as to whether OSG's have any benefit or indeed have harmful consequences for users impedes their use by organisations who require evidence to make informed decisions as to whether they are suitable to use.

However, Barak et al. (2008) offer another perspective on the efficacy of OSG's based on a review of qualitative studies of OSG's. They argue that the primary benefit of OSG's is in fostering a sense of personal empowerment (well-being, a sense of control, self-confidence, independence, social interactions, and improved feelings). Several factors within OSG's combine together to produce this effect, including the disinhibition effect (accelerates personal and interpersonal dynamics), impact of writing, expression of emotions, improving knowledge and understanding, developing relationships, improving decision making and as a result behaviours. Barak et al. (2008) also highlight the pitfalls of OSG's such as becoming overly dependent on the group, becoming more removed from face-to-face contacts and having negative or unpleasant experiences peculiar to online interactions. Subsequent studies have shown that OSG's do appear to foster a sense of personal empowerment (Mo & Coulson 2010; van Uden-Kraan et al., 2008).

Extending on the theory that OSG's foster a sense of personal empowerment, Lawlor and Kirakowski (2014) argued that OSG's may reduce self-stigma (internalised stigma) through personal empowerment among those with mental health issues, which

in turn would encourage help seeking offline. This hypothesis is based on the premise that peer to peer contact helps to create a positive group identity that challenges self-stigma to create a positive self-identity (Corrigan et al., 2009). As self-stigma is the biggest barrier to seeking help for mental health issues, removing this barrier through OSG participation should increase help seeking behaviour (Corrigan et al., 2011). The results from Lawlor and Kirakowski's (2014) study did not wholly support the hypothesis, in fact the reverse was found to be true. Frequency of visits negatively affected recovery from self-stigma which indicated that OSG's were functioning as a source of social avoidance rather than encouraging social participation. Users who actively participated did have higher levels of recovery from self-stigma and were more likely to have sought formal support. However, this could be attributed to underlying factors which are associated with active participation rather than because of active participation.

The anomaly in the findings, that is frequency of visits negatively affected recovery from self-stigma, was primarily suggested to be explainable through social avoidance. However, Lawlor and Kirakowski (2014) also offer alternative theories that may account for the anomaly in the findings. One such theory was the possible presence of MbI perpetrators within the OSG's that the participants were recruited from. The rationale for the MbI theory was based on mounting evidence that MbI perpetrators were a feature of OSG's and that their unwitting inclusion in data analysis needed to be acknowledged by researchers. Anecdotally, cases of MbI were being increasingly reported in the media, particularly within OSG's for eating disorders, cancer, and pregnancy (Kleeman, 2011; Moran, 2012; Jenkins, 2012; Szalavitz, 2012). In response to the increased perceived threat from MbI perpetrators, online users were mobilising themselves into groups internally and externally of OSGs to investigate suspected cases of MbI e.g. warriorelihoax.com, fake-lj-deaths.livejournal.com. Researchers also became interested in MbI, since the very first case studies conducted by Feldman, Bibby and Crites (1998), another twenty case studies of MbI have been unearthed (Cunningham & Feldman, 2011; Witney, Hendricks & Cope, 2014; McCulloch & Feldman, 2011; Cadet & Feldman, 2012; Feldman, 2000; Uridge, Rodan & Green, 2012; Vanelli, 2002; Ayyer & De Sousa, 2014; Feldman, 2004). Of these twenty-four case studies, two were encountered as part of wider research projects that had created OSGs that were infiltrated by an MbI perpetrator (Witney, Hendricks & Cope, 2014;

Uridge, Rodan & Green, 2012). These studies helped to underscore the importance of researchers being cognizant of MbI perpetrators and factoring in this variable during analysis. Additionally, Eysenbach and Till (2001) highlighted the risk of the incorrect interpretation of data when false stories of personal illness are unwittingly included. Researcher acknowledgment of MbI is especially important because the prevalence is assumed to be high as the internet offers an unrivalled opportunity for impression management and innumerable groups to solicit attention and sympathy (Pulman & Taylor, 2012).

Considering the mounting evidence that MbI is an extraneous variable that needs to be factored into data analysis, Lawlor and Kirakowski (2014) examined the anomaly in the findings from this new perspective. That is the possibility that MbI perpetrators were among the participants and the secondary effect that their presence would have on the genuine participants who interact with them, knowingly or unknowingly. Investigating the anomaly in the findings considering MbI was further supported by overt incidences of MbI being retrospectively observed by Lawlor and Kirakowski (2014) in some of the OSGs that participants were recruited from. Directly, with regards to the veracity of the participants, if MbI perpetrators were included as participants a negative effect on recovery measures would be expected. This is because the modus operandi of MbI perpetrators is to obtain sympathy and attention that is elicited through their circumstances worsening (Feldman, 2000). Furthermore, there was no significant difference between the overall mental health scores of participants who claimed to have or had not received informal support and formal support. In a 'normal' population this result would be unexpected and may be further evidence of MbI where the illness is maintained for attention and sympathy. Indirectly, the mere presence of MbI within an OSG may have an adverse effect on genuine users by populating the OSG with misinformation about illness progression and prognosis, and cultivating a general negative attitude within the OSG (Pulman & Taylor, 2012; Witney, Hendricks & Cope, 2014). A negative attitude instigated by the presence of MbI could also account for the lack of recovery from self-stigma reported by participants. This is because recovery requires the formation of a positive self-identity through the positive group identity (Corrigan et al., 2009). However, an OSG frequented by MbI perpetrators will create the opposite environment, one dominated by negativity and hopelessness, so they can receive as much attention and sympathy

as possible. This negativity may be absorbed by genuine users hindering their ability to recover from self-stigma.

Ultimately these were just theories put forward to try and offer explanations for the results. As there was no previous literature on the potential of MbI to negatively impact genuine users or a method of detecting and thus removing these individuals, no definitive conclusions could be drawn. This gap in the literature was the impetus for this thesis. The objective therefore was to develop a method of detecting MbI. In choosing an appropriate method of detection more had to be known about the first-hand experiences of those who suffer from FD and online users who encounter MbI. The findings from the studies presented in this thesis were not just confined to developing the classifier to detect MbI, but also had wider theoretical implications regarding our understanding of online communication and deception. Before undertaking these studies, a literature review of FD, MbI and MbI's position within digital deception literature was conducted to give direction and wider context to the research and this is presented below.

1.2 FACTITIOUS DISORDER

1.2.1 DIAGNOSIS

Factitious Disorder (FD) (commonly referred to as Munchausen's) is a psychological disorder, where a person pretends to be physically and/or psychologically ill or tells elaborate stories to elicit attention and affection, by occupying the sick role. The DSM-V requires that the following criteria are met before making an official diagnosis of FD:

A. Making up physical or psychological signs or symptoms or causing injury or disease with the deliberate intention to deceive
B. Pretending to be sick or injured or to be having problems functioning
C. Continuing with the deception, even without receiving any visible benefit or reward
D. Behaviour is not better explained by another mental disorder, such as a delusional disorder or another psychotic disorder

Within this category there are two variants of FD which are recognised by the DSM-V:

(1) FD imposed on self
(2) FD imposed on another (previously FD by proxy)

The previous diagnostic criteria provided by the DSM-IV were widely criticised and despite the revisions made to the DSM-V the underlying issue, that research does not support the criteria persists. This is due to the difficulty in accessing people with FD. Consequently, the DSM criteria are vague and do not describe the symptoms that the sufferer is experiencing, but rather the effect it has on those around them (Kanaan & Wessely, 2010). Hamilton et al. (2009) highlight that FD is not compatible with the underlying assumptions of the DSM. The first assumption is that the disorder should be distinguishable from other disorders. However, FD is identical to both Somatoform Disorder (involuntarily produced symptoms) and Malingering (voluntarily produced symptoms to gain external benefits) when only observable behaviour is considered. To distinguish FD from these disorders, the individual's internal motivation for assuming the sick role must be known, but this is impossible to ascertain from observations. The second assumption is that the disorder is discrete, you either have it or you don't. However, it is more likely to vary in severity. For example, Reich and Gottfried (1983) estimated that 10% have Munchausen's, a severe chronic form of FD. The third assumption is that the individual suffering from the disorder should be able to self-identify. Those suffering from FD tend not to want to be identified as they are reliant on the sick role to meet their emotional needs and they may feel ashamed of their behaviour, even though they are suffering from a genuine psychological disorder. In practice, the diagnosis of FD using the DSM is significantly lower than one would expect. This is due to clinicians not wanting to alienate their patients and the difficulty involved in ascertaining the suspected patient's motivation for assuming the sick role, that is, how can intentionality be determined (Kanaan & Wessely, 2010). Therefore, conclusive diagnosis of FD is not dependent on the DSM criteria, but rather on the patient either admitting to feigning or being caught in the act of trying to produce

symptoms. Bass and Haligan (2007) argue that in most cases FD is diagnosed because the sufferer has a history of feigning illnesses.

1.2.2 SYMPTOMS

Due to the lack of first-hand accounts from those suffering from FD, the internal symptoms that are experienced are largely unknown. However, Catalina et al. (2008) created a list of 8 observable behaviours associated with enacting FD. The simultaneous presence of three of these behaviours is believed to be indicative of FD:

- Inconsistent response to treatment.
- Inconsistent symptoms.
- Symptoms worsen after discharge.
- Symptoms disappear after admission.
- Develop intense relationships with staff and other patients.
- Symptoms are like other patients.
- Lies.
- No evidence of disorder claimed.

1.2.3 PREVALENCE

As diagnosing FD is problematic, the prevalence is difficult to determine, but it does appear to be dependent on the type of illness that is being feigned, as some are more amenable to feigning than others. There are three criteria described by Kanaan and Wessely (2010) which make an illness amenable to enacting FD:

- Readily available knowledge of the illness.
- The symptoms are easily reproduced.
- Illnesses which elicit the most amount of sympathy are the most desirable.

Ishak et al. (2010) reports that studies have estimated prevalence rates of between 0.5% and 0.8% in hospital patients and 6% to 8% in psychiatric patients. Due to the diagnostic problems and the reluctance of clinicians to diagnose FD, the real figure is likely to be higher. One of the highest prevalence rates of FD has been for neurological and dermatological problems, at around 15%, as the symptoms are easy to feign.

1.2.4 CHARACTERISTICS

With regards to FD, Pridmore (2006) highlights that most sufferers are women (72%) who are working in the health profession (66%). They tend not to present to multiple treatment centres, have a history of not having their emotional needs met and are experiencing sexual and/or emotional problems. In comparison, those with the severe form of FD, Munchausen Syndrome, are typically single males in their 40's who have a Cluster B Personality Disorder (Anti-Social, Borderline, Histrionic and Narcissistic Disorder). They travel to multiple treatment centres because once confronted they tend to leave the treatment centre and move on to the next. The lies they tell also tend to be extraordinary (Pridmore, 2006). There are also some characteristics associated with FD more generally, including having a Personality Disorder, being raised in an institution, enjoying being the centre of attention, the feigned condition becoming central to their life and feeling superior to professionals as well as having a medical background (Hallstrom, 2009).

1.2.5 MOTIVATION

Understanding why FD sufferers pretend to have physical/psychological disorders is pivotal to developing effective treatments and diagnostic methods. However, theories in relation to motivation have been purely speculative due to lack of first-hand accounts from sufferers. The following are some of the various theories that have been put forward as possible motivations for falsely assuming the sick role (Ford, 1996; Dryer & Feldeman, 2006):

- Gratification of dependency needs – Searching for nurturing.
- Defence against psychosis – Lack the ability to cope, as such become prone to acute anxiety which can become overwhelming causing a decline in reality testing. They create a role for themselves which gives a sense of identity. They are simultaneously in control (altering behaviour to feign illness) and dependent (assuming the sick role).
- Need for identity – They have a poor sense of self and through the sick role they become to feel important.

- Need for mastery – By ‘fooling’ others around them into believing they are sick, they feel clever and important and experience pleasure i.e. ‘duping delight’.
- Internalised anger/masochism – They involve themselves in self-defeating behaviour, believing that they deserve abuse they seek nurturing through pain.
- Learning theory – Earlier in life they learned that their emotional needs were met when they were sick. Therefore, they assume the sick role in order to replicate this scenario later in life.
- Psychodynamic Theory – They had a problematic child/parent relationship which resulted in the need to assume the sick role.
- Need for distraction from authentic life stressors – By feigning an illness they can mask the real problems they are experiencing and therefore do not have to confront their real issues.
- Self-Enhancement Model – They feel special by being ill and having the attention of high status individuals e.g. doctors.

1.2.6 TREATMENT

Treatment centres on understanding the emotional life of the FD sufferer and developing coping strategies and more appropriate ways of communicating emotional needs (Pridmore, 2006). There is very little robust research which examines the effectiveness of treatments. A systematic review by Eastwood and Bisson (2008) found no difference in the outcome of treatment with regards to directly confronting the sufferer with their behaviour or not, and with regards to whether psychotherapy was received or not and whether medication was prescribed or not. It was noted that there was a problem with keeping FD sufferers engaged in treatment and that long-term management strategies needed to be developed. This difficulty in keeping sufferers engaged in treatment has been noted by other authors, because sufferers are also reluctant to seek help to begin with, partly because they do not acknowledge that they are enacting FD. For example, a study by Krahn et al. (2003) reported that among 93 hospitalised patients with FD, 17.2% admitted that they had feigned the disorder and 11.8% of inpatients and 8.6% of outpatients agreed to seek help. Recovery requires admitting to the disorder and then engaging with long-term therapy, which is unfortunately uncommon.

1.2.7 FACTITIOUS DISORDER IMPOSED ON ANOTHER (MUNCHAUSEN BY PROXY)

1.2.7.1 DIAGNOSIS

FD imposed on another (FDIA) is a unique subtype of FD. It is not recognised as a mental disorder but rather a form of abuse where a caregiver falsifies symptoms of a psychological/physical illness in the person under their care so they can vicariously assume the sick role. We distinguish here between the perpetrator (the person carrying out the falsehood) and the victim (the person about whom the false allegations are made) although "perpetrator" may be too judgemental a term for what is a mental disorder. The DSM-V requires the following criteria are met before making a diagnosis of FDIA:

A. Making up physical or psychological signs or symptoms or causing injury or disease in another person with the intention to deceive
B. Presenting another person to others as sick, injured or having problems functioning
C. Continuing with the deception, even without receiving any visible benefit or reward
D. Another mental disorder does not better explain behaviour

As with FD, FDIA diagnosis is confronted with the same issues, how to determine intention and motivation which are internal states. Instead numerous studies have looked at what observable behaviours are associated with the perpetrator to create a suspicion index. However, Pankratz (2006) reports that by 1996 there were 100 warning signs and that many of these contradicted one another (calm/distressed, charming/hostile, and distant/over involved) and could easily be applied to innocent care givers. Not all warning signs are centred on the perpetrator and they also include characteristics of the falsified illness including the illness being multi-systemic, prolonged, rare, symptoms being inappropriate or incongruent, the patient having multiple allergies, having poor tolerance of treatment and seizure activity which is unresponsive to treatment (Khan, 2008).

Once suspicion has been raised, an attempt to confirm the diagnosis will made. This is usually done by using lab tests, collecting information from relatives and ascertaining if the patient's symptoms disappear when the caregiver is absent (First & Tasman, 2010). In some instances, video surveillance may be used, however Pankratz (2006) argues that describing what is happening on these tapes is not clear cut. While these practices do provide evidence of abuse, the question remains as to whether they diagnose FDIA. This has led to the diagnosis being highly controversial and Long (2008) argues that at the root of the controversy is the merging of a criminal act and mental disorder. As a mental disorder, it is poorly defined and there is confusion over who diagnoses it and how. This has led to the view that falsification of symptoms in another should be looked on as an instance of abuse. Then if necessary a psychologist/psychiatrist will offer an opinion on the perpetrators mental state, but not as to whether the perpetrator has FDIA because at present there is no scientifically validated method of making such a diagnosis. This is not to say that FDIA is not a mental disorder. This type of abuse does differ from typical abuse, as the motivation is not to exercise frustrations, but to have emotional and social needs met (Criddle, 2010). Therefore, it is the responsibility of psychologists and psychiatrists to develop a more sophisticated method of diagnosing FDPB.

1.2.7.2 PRESENTATION AND PREVALENCE

Although not officially in the DSM-V, Schreier and Libow (1993) describe four types of FDIA:

- Chronic – Constant falsification of symptoms in another to the point where it becomes the central focus of their lives.
- Episodic – Periods of remission where they do not feel the need to falsify symptoms in another.
- Mild – Creating a fake medical history or pretending the person has been sick, as opposed to inducing symptoms.
- Intense – Physically harming another by inducing symptoms e.g. removal of blood, administering harmful substances, suffocation etc.

A review by Sheridan (2003) found that the most common symptoms which are falsified in another include apnoea, anorexia/eating problems, diarrhoea, seizures, cyanosis, behavioural problems, asthma, fever, allergy, and pain not specified. In approximately 60% of cases the symptoms were induced through suffocation, drugs, and poisoning. On average, it takes 22 months between onset of symptoms and diagnosis of FDIA. The prevalence of FDIA is difficult to estimate due to the diagnosis issues previously described and is likely to vary by the type of illness feigned. McClure et al. (1996) calculated that the incident rate in Ireland and the UK for children under 16-year olds was 0.5/100,000 and for under 1-year olds, 2.8/100,000.

1.2.7.3 CHARACTERISTICS

In the case of victims, Sheridan (2003) reported that if the mother is the perpetrator the victims are as likely to be male as female. If the father is the perpetrator, they are more likely to abuse their son. The average age at diagnosis is 48 months. With regards to perpetrators, 75% are the victim's mother followed by fathers at 7%. Approximately 15% had experience in the medical profession, mainly as nurses. 30% displayed characteristics indicative of Munchausen Syndrome. Of the 23% with a psychiatric diagnosis, the most common were depression, personality disorder and paranoia. 22% experienced abuse either during childhood or by a partner.

1.2.7.4 MOTIVATION

The motivation for FDIA is likely to vary according to the emotional and/or social needs the perpetrator is inappropriately trying to fulfil. Criddle (2010) lists several theories that have been put forward:

- Receiving sympathy, attention, respect, and public acclaim by appearing to be devoted.
- Feeling in control by demonstrating their perceived superior knowledge by manipulating medical professionals.
- Avoiding daily responsibilities.
- Providing themselves with a purpose, making them feel important and interesting.

- Seeking revenge on a spouse.
- Receiving support from friends and family.
- Using the hospital as a social life.

1.2.7.5 TREATMENT

The mortality rate for victims of FDIA is high, 6% to 9%, with over half related to apnoea (First & Tasman, 2010). The long-term outcome for surviving victims is variable depending on the initial treatment received. There may be permanent disability, continued falsification of symptoms, somatoform disorder, emotional disorders, conduct problems and problems in school (Dingle, 2005). 8% will have chronic physical problems and the majority will experience psychological symptoms (Criddle, 2010). Therefore, it is important for victims to seek therapy as soon as possible. For the perpetrators, if the abuse is not detected and treatment is not received, the chances of the perpetrator abusing again is high. For example, Sheridan (2003) reported that 61% of the siblings of victims of FDIA had also been victimised and 25% had died suddenly. Treatment primarily requires the perpetrator to acknowledge what they have done and that they need therapy. Therapy then focuses on identifying what is perpetuating the disorder and how this can be managed. Family therapy can be used to facilitate the possible reunification of the family with supervision. Those with a good chance of recovery take responsibility for the abuse they have inflicted, develop coping mechanisms, and put their children's needs above their own (Brannon, 2015).

1.2.8 IMPACT ON THOSE DUPED BY FD AND FDIA VICTIMS

Both FD and FDIA have a significant negative impact on those duped. The most immediate victims of FDIA are those of by Proxy who have been abused. However, FD also has a significant negative effect upon those who have close contact with sufferers including health professionals, family, and friends. Criddle (2010) argues that one of the primary challenges faced by health care professionals is that they have unwittingly been involved in putting a FDIA victim or FD sufferer through numerous unnecessary invasive treatments, which can contribute to morbidity. 75% of morbidity in FDIA victims is due to multiple invasive procedures. There are several theories as to why health care professionals fail to consider FD or FDIA and instead consider the possibility of a statistically less likely rare disease:

- They are very convincing.
- Health care professionals are interested in difficult and challenging cases.
- Some health care professionals maybe too thorough when trying to make a diagnosis.
- They may ‘doctor shop’ for doctors who meet the above two criteria.
- They could threaten litigation by claiming the physician has not done enough.
- General disbelief that anyone would behave in this way.

Maldanado (2002) reports that FD and FDIA can also divide the health care team into two opposite groups, those who believe and those who do not, causing significant friction within the team. When FD or FDIA is diagnosed, there is a feeling of betrayal, anger, resentment, and an increased suspicion of people faking illnesses in themselves and others. Therefore, it is important for the team involved to receive some form of therapy to work through these residual issues.

Perlmutter (2006) describes the impact on friends and family as a sequence of conflicting emotions. Before a diagnosis of FD or FDIA, it is likely that family and friends will have gone through a period of intense worry about the severity of symptoms, dramatic course of the illness and the fear that the ill person will die. Inevitably, treatments will not work resulting in a cycle of anger towards the person for not getting better and then guilt for being angry at a sick person in the first instance. Anger will then be focused towards health care professionals for their inability to make a diagnosis and due to stories of ineptness created by the FD sufferer or FDIA perpetrator. After the mix of emotions that family and friends have been put through, the diagnosis of FD is met with confusion (“how could anyone do this?”), anger (“what are you trying to pull?”), and denial (“Let’s leave, they clearly don’t know what they’re doing”). There can be a movement from denial to anger, rage, disgust, and a sense of betrayal towards the FD sufferer or FDIA perpetrator. Therefore, therapy for all those involved is so important because this disorder does not just affect the sufferer.

1.2.9 STIGMA SURROUNDING FD

Pridmore (2006) has highlighted how FD is subject to the stigma which surrounds mental health disorders. But this stigma is further compounded as sufferers of FD are viewed as wasting time and resources which could be channelled towards those who are genuinely ill. However, they are genuinely ill and as such they are entitled to occupy the sick role. Unfortunately, they do not know how to communicate their actual problems or may not be aware that they have a legitimate illness, FD. Pridmore (2006) has put forward three theories to explain the stigma which surrounds FD. Firstly, at the root of the stigma is the perception that FD sufferers violate the social norms of being sick, the obligations and privileges of occupying the sick role. These obligations include, the perception that occupying the sick role is undesirable, the need to co-operate with others (doctors, nurses etc.) in order to become well again and making use of the appropriate services which are available to aid diagnosis The privileges of the sick role include, having no responsibility for their state as they have not purposely caused it, being recognised as someone who requires special care and being entitled to exemption from regular obligations, for example, work, school and daily chores. Those with FD are viewed as breaking these social norms as they have access to the privileges of the sick role without fulfilling their obligations i.e. they appear to enjoy the sick role. They are not co-operating to become well again (lying, feigning symptoms, aim is not to become well) and they either overuse (for their feigned illness) or underuse (for their real disorder, FD) available services. Therefore, they are viewed as gaining unfair privileges. Secondly, Pridmore (2006) argues that the issue as to whether FD sufferers deserve the sick role or not is purely a western one. This is because in the West the dominant philosophy is dualism (mind and body are separate) as opposed to monism in the east (mind and body are inseparable). Through a monist perspective, being ill is applicable to the whole person and not just isolated to the mind or body. Therefore, if an individual is feigning symptoms of an illness, this is a symptom of their greater illness (FD) and so they are entitled to the sick role. It is not just a simple case of lying. Thirdly, from an evolutionary standpoint, the concept of 'reciprocal altruism' may partially explain the stigma. This concept relates to the idea that there is safety in numbers, so to maintain the group number, help will be given to another in need if it will not significantly cost the helper. An individual who abuses

this system (feigns illness) will be looked on negatively and excluded, as they will have unnecessarily cost the helper.

The consequence of this stigma is reluctance to seek help, Hagglund (2009) describes the process by which stigma can discourage an individual with FD from seeking help. Firstly, there may be a reluctance to officially diagnose FD by the clinician in case the sufferer will be met with scepticism from health care providers in the future. The sufferer themselves may avoid being diagnosed for fear of the reaction from family and friends, who may have been negatively affected by their behaviour, for example, worrying about a loved one's health and then finding out it was feigned. Secondly, they may believe that being physically ill is more 'legitimate' than being psychological ill, due to the public stigma associated with mental health disorders. Therefore, they are going to be less likely to want to adopt a diagnosis of being mentally ill. Thirdly, family members may be subjected to courtesy stigma, whereby the stigma of being mentally ill, is not just confined to the person suffering from a mental disorder, but is also applied to anyone related to that person. As such, family members may unconsciously react more compassionately towards physical symptoms as opposed to psychological, reinforcing the idea that a feigned physical complaint is more desirable than a legitimate diagnosis of FD.

1.3 MUNCHAUSEN BY INTERNET

MbI is a variant of FD whereby a person uses the anonymity of the internet to masquerade themselves as being mentally or physically ill, a victim and/or hero using elaborate fictional stories. MbI is not officially recognised by the DSM-V, however calls have been made for its inclusion within the FD category (Pulman & Taylor, 2012).

1.3.1 CASE STUDY OF MBI

The best way of understanding MbI is through a real-world example. A case study is provided to help illustrate MbI and quotes from users who were duped are also included. These quotes help to illustrate the extent to which people can become emotionally invested in MbI perpetrators. Below is a synopsis, that has been created

for this study, that recounts a post made by a web bloggers wife who logged into a blogging site to tell members that her husband had committed suicide.

In her post, the web blogger's wife tells the community how much her husband loved the group, how they joked about the group being his 'mistress' and the invaluable support and advice he had received down through the years. She recounts the numerous issues and difficult life circumstances he had faced and openly shared with the online community. He was a former methadone addict and was sexually abused by his Uncle. His mother was a lifelong addict arrested for child endangerment and died of AIDS. He had a 13-year-old son who he had become estranged from due to his addiction but they were making progress towards rebuilding the relationship. He had just found out he had a 14-year-old daughter and was slowly building a relationship with her. His father and step mother were emotionally abusive. He suffered from crippling social anxiety. Recently he felt partially responsible for the murder of a girl because he provided a statement which led to the release of her boyfriend who subsequently murdered her. His wife talks about how happy they were together and their little idiosyncrasies. Saying 'leprechaun, leprechaun' meaning 'I love you'. How she could guess his mood by asking what music he was listening to. How he wrote a comic strip for her when he was in jail and the motivational and spiritual postcards he would send to her. She goes on to describe in detail the events leading up to his suicide. She went to her parents for the weekend. Her husband did not go as he had previously had a conversation with her father about God, science and religion which had led to her father questioning his faith. This in turn did not please her mother, so he decided to stay home. The next morning, he was due to meet a friend to play disc golf, his friend heard the car engine was running inside the garage and found her husband had died from fume inhalation. Initially she believed he had not planned the suicide in advance, even questioning if it was a murder made to look like suicide. However, a few days later she received one of the postcards he liked to send her. He explained that he was struggling with his methadone addiction and had always said he would die before he went back on it again.

Underneath this post there are 682 comments from members of the weblog describing how shocked and stunned they are by the news and offering their condolences, below are some examples:

I've just now seen this and I am typing through my tears. I am very, very sorry for your loss.

His comments and posts always made me look at life in a brighter light, even if they dealt with dark issues. Thanks for sharing your story. I'll miss his presence and will never forget his openness and kindness.

I am so terribly sorry for your loss. I would say "may his memory be a blessing to you," but it is you who have blessed us with the memories you have shared. You and he and everyone who loved him will be in my thoughts.

What you have shared here is profound, as was your husband's life. It feels like a true privilege that we've all had the chance to share in it. I want you to know that part of my heart that has been numb for a very long time was opened up and lit aflame by reading and re-reading your and your husband's words all day today. I will not forget them.

Six days after the post that her husband had committed suicide, the moderators made the announcement that they discovered it had been a hoax. They had become suspicious as there was no obituary notice, no condolences on social media and there were references to a friend socializing with the weblogger a few days after the claimed suicide. When confronted, the weblogger admitted that it was a hoax. The response to the exposure was one of incredulity and anger, but some were philosophical about the whole experience stating that the outpouring of support towards the weblogger's wife only demonstrated the goodness of the community. Below are some examples of the comments:

This is seriously fucked up. I feel so disgusted knowing that this happened. Someone tried to take advantage of the community and that's really upsetting. I don't know what to say.

The takeaway I got from that thread is that this is an amazing community full of people who have mostly never met, but still have nothing but love and support

for each other.

I was momentarily angry, but I think I'm mostly sad, because, you know, there are a lot of people going through a lot of shit, and it's hard to reach out, and now some of them who do take that really hard step, I won't believe or I'll be sceptical about because I was deceived in this fashion. I'll be a slightly less generous and more cynical person.

This case study of online pseudocide, involving the use of a sockpuppet, (alternate personalities used for deception) helps to illustrate the stages of an online hoax, the detailed nature of the ruse, the genuine fondness members have towards the perpetrator before the hoax is revealed and the mixed reactions they have when they realise they have been duped. It is a typical case of MbI that is not an uncommon occurrence in online communities.

1.3.2 Overview of MbI

MbI was first identified by Dr Marc Feldman in 2000, he described it as the misuse of online communities by offering false stories of personal illness and crisis to garner attention and nurturing. MbI is the online expression of FD, however it is not recognised by the DSM-V. A non-pathological version of MbI also occurs online, malingering. Eisendrath (1996) describe malingering occurring when there is no internal incentive for the behaviour only external i.e. soliciting money and gifts. FDIA is also expressed online and is referred to as Munchausen by proxy by Internet (MBPBI) (McCulloch & Feldman, 2011). MBPBI typically involves pretending a relative or close friend is sick and posting about the experience online on their behalf, as a supporter and/or creating a sockpuppet representing the sick person. The final type of online deceptive behaviour that involves occupying the sick role that was proposed by Pulman and Taylor (2012), MbI by Trolls (MBIBT). This is a more sinister form where the aim is to disrupt and abuse the online group for enjoyment.

As seen in the case study, these ruses are not just confined to trying to occupy the sick role, they also involve a lot of additional embellishment. This is referred to as "pseudologia fantastica" (Sadock & Sadock, 2011; Feldman & Ala, 2000). The use of

"pseudologia fantastica" helps to create a captivating and compelling story that catches the reader's attention which in turn feeds into the MbI sufferers need for attention and sympathy. King and Ford (1988) describe the four characteristics of "pseudologia fantastica": (1) the stories are not completely improbable and are frequently grounded in a matrix of truth; (2) the stories are enduring; (3) the stories are self-aggrandizing; (4) when confronted with evidence the perpetrator can acknowledge the falsehoods. Another important characteristic noted by Dupre (1909) is that while the themes of the stories may be varied, they will most certainly revolve around the hero/victim role. Therefore, MbI is not just limited to occupying the sick role but also can involve the occupation of the hero and/or victim role.

1.3.3 MOTIVATION

The motivation for engaging in the four-online deceptive behaviours (MbI, MBPBI, MBIBT and malingering) are based on a very limited number of first-hand accounts from those who confess. Therefore, theories regarding motivation are mainly dependent on inferences from what is known about their offline equivalents, which are speculative to begin with. The following sections will discuss what is currently known about each of the online deceptive behaviours based on extrapolations from FD and the limited first-hand accounts that are available.

1.3.3.1 MBI

Ayyer and De Sousa (2014) offer a rare insight into MbI sufferers through the case study of an 18-year-old male who was brought for a psychiatric consultation by his parents. This was prompted after his deceptive behaviour was accidentally exposed. A member of an online support group for cancer contacted the perpetrator's family trying to encourage him to go to the support group gathering in his town. It then emerged that he had been posting regularly on an online cancer support group where he falsely claimed to have terminal stomach cancer. He had even befriended one member who was also terminally ill and they had agreed to offer each other companionship. When confronted by his parents he was callous, mocking members for their naivety and alluding to the incident as a mere prank. However, after further investigation it emerged that the motivation for the behaviour was deeply rooted, as seen from the extract below (Ayyer & De Sousa, 2014, p. 82):

He felt that his parents did not care for him or understand him and he could not share his feelings with others. He believed ill people got a lot of care and attention from others and lying on internet sites made him feel wanted. He believed it was a harmless way of getting attention and fantasized about group members replacing his family. He had no insight that this was a psychological problem.

Similarly, Swains (2009) reports on a rare case described by Feldman where the perpetrator had also tried to explain her motivation (para. 7):

I have never felt more loved and cared for in my entire life. I suddenly craved for everyone's attention, love, care, concern and affection…It became very appealing to me. I decided to play with it more. I do not know how or why, I just did.

The motivation for MbI expressed in both cases was to assume the sick role for its psychological benefits. This is also speculated to be the primary motivation for MbI's offline equivalent FD. Tasman and Mohr (2011) state that the need to occupy the sick role for these psychological benefits is because of pre-disposing factors that result in psychological deficits. The pre-disposing factors are thought to be rooted in childhood trauma/neglect which in some instances can also result in the development of personality disorders. There is a strong co-morbidity between FD and Borderline Personality Disorder (Sadock & Sadock, 2008). There are a range of psychological deficits believed to be associated with FD, but they are largely unsubstantiated (Ford, 1996). The psychological deficits most likely associated with enacting MbI include people with a poor sense of self using the sick role as a structure to develop their identity around, having unmet dependency needs met through the attention, caring and nurturing associated with the sick role and, for those who have experienced trauma, (childhood illness, abuse) creating a fictional trauma as a way of getting back control and mastering their real trauma (Fleming & Eisendrath, 2011).

1.3.3.2 MBPBI

Anecdotally speaking, just as with FDIA, the most common type of MBPBI involves a parent, typically mother, pretending they have a child who is sick. They blog about

their experiences and/or join an online community and start posting about their tribulations. It's important to note that often the deception is not exclusively online and can be accompanied with offline deception. There have been cases where a mother was posting online about their sick child in addition to intentionally poisoning their child offline (Pavia, 2015). The highest profile and intricate cases of MBPBI, that included a first-hand account from the perpetrator as to their motivation, was the 'Warrior Eli blog'. The ruse ran for almost a decade. The fake blog was originally set up for the son of John and Emily Dirr, to chronicle his battle with childhood cancer. It was given credibility through the creation of 71 sockpuppets in the form of friends and relatives who communicated back and forth with the Dirrs. Suspicions were raised due to the series of tragedies which befell the family, culminating in Emily Dirr's death in a car crash on Mother's Day. She miraculously managed to stay alive long enough to give birth to the families eleventh child. The hoax was exposed with the help of a blog set up by Wright (2013), when the pictures of the Dirrs and their various friends and relatives were found to be stolen from other people's social media accounts. Once confronted, the perpetrator admitted to the hoax and was revealed to be a 22-year-old female medical student training to be a doctor. She offered this apology via Wright's blog (Wright, 2012, para. 3):

> *To whom I have hurt: I am deeply sorry for all the pain I have caused everyone. It was never my intention to do so. This all started 11 years ago when I was a bored 11-year-old kid looking for an escape from the pain and heartache I saw in my own family. It started almost as a fiction writing, but the more time I spent escaping to it, the more "real" it became. I am so sorry it hurt so many real families, and so many people out there.*

Apart from these rare first-hand accounts of the motivation behind MBPBI, there is little known about the disorder from the perpetrator's perspective. Furthermore, the sincerity and accuracy of such confessions and apologies also must be met with scepticism. For example, in exposing the 'Eli Warrior blog' hoax, Wright (2012) stated that the perpetrator in their initial confession had told them a sob story which turned out to be untrue. As there has been no independent research on MBPBI, theories regarding motivation are inferred from FDIA. Those who perpetrate FD imposed on another do so to fill a psychological deficit, as is the case with FD. They seek to attract

attention, sympathy, care, concern, and a sense of importance from being the parent of a sick child as well as to manipulate and control others (Lasher & Sheridan, 2004).

1.3.3.3 MBIBT

Pulman and Taylor (2012) have argued that there is a need for a new category, MBIBT, where the motivation is primarily rooted in a co-morbid personality disorder. It is purely for enjoyments sake to feed sadistic and/or narcissistic personality tendencies. Swains (2009) reports that Feldman also expresses the belief that there is a more sinister subset of MbI involving a pre-existing personality disorder which primarily involves the manipulation of online communities and their members. Feldman argues that in some cases of MbI, the feigned scenarios tend to be highly emotive and are followed by a superficial apology when exposed, suggesting an element of sadism. There is also a sense of being 'in control' by controlling the thoughts and reactions of others online. Similarly, Eisendrath (1996) notes that FD is linked with co-morbid personality disorders, namely Borderline Personality Disorder. The borderline personality is characterised by unstable interpersonal relationships, rage, impulsiveness, self-harm/suicide attempts, rapid changes in mood and switching between idealizing and belittling others. When combined with FD, duping others allows them to focus their anger thus temporarily stabilizing their mood. Duping online community members by falsely occupying the sick, victim and/or hero roles could be another outlet for those with a co-morbid Borderline Personality Disorder to focus their rage resulting in MBIBT.

1.3.3.4 MALINGERING

Malingering is distinct from the other types of deceptive online behaviour involving the occupation of the sick, victim and/or hero role in that the motivation is purely for external incentives (Hamilton, Feldman & Janata, 2009). Anecdotally, the most common type of ruse seems to involve an element of by Proxy (Monroe, 2016). The online malingerer creates an emotive story by pretending their child is ill. They then extort money from members of the online community by claiming it is required for treatments, travel expenses or gifts for the child. The distinction between malingering and FD is not however always clear cut. This is highlighted by Worley, Feldman and Hamilton (2009) who argue that there is a difference between motivation and consequence. Therefore, a person with FD might find themselves in receipt of external

incentives that accompany the sick role, but this was not their primary motivation. This too applies to MbI where well-meaning community member may offer material support as well as emotional.

1.3.4 Issues with applying extrapolations from FD to MbI

At this juncture, it's important to address the problem of applying extrapolations from FD to MbI. Firstly, there is a dearth of information available about FD given the elusiveness of sufferers and what is known is based on inferences from observational case studies. Secondly, Bass and Haligan (2007) have argued that the fixation on diagnosing and distinguishing FD based on motivation is inherently problematic as motivation is highly variable and impossible to establish. Thirdly, Ford (1996) notes that there is a multitude of motivations put forward to explain FD and they are unsubstantiated. Instead Bass and Haligan (2007) believe the focus should be on the underlying psychiatric problems which precipitated the deceptive behaviour. Therefore, our understanding of MbI is ultimately restricted by our limited understanding of FD.

1.3.5 Impact of MbI on those duped

One of the most dangerous consequences of MbI, is that many of their fictitious stories include false information about personal history, medical advice they have received, treatments, medical tests, and information from medical publications as part of their ruse. Whitney, Hendricks, and Cope (2015) believe that this misinformation could then be used by those who read it to make misinformed decisions about their own health. Mo and Coulson (2014) note that the experiences shared in online support groups have been shown to be used to inform health care and disease management decisions. However, despite the potential fatal risks, only one study has been conducted to examine the effect of MbI on online communities and their members. Feldman (2000) reported the effect that exposing hoaxes had in four case studies of MbI, the reactions included:

- Division of community into those who believe and those who do not.
- Staying in community to deal with feelings of anger, sadness, or shame.
- Leaving the community in revulsion.

- Communicating feelings to the deceiver through email.
- Seeking revenge by exposing deceiver in their daily lives.
- Fantasising about directly confronting the deceiver.
- Fearing deceiver will misuse personal information they have exchanged with them.
- Finding the deception amusing.

While these do illuminate some of the superficial effects of MbI they do not provide any deep insight. Feldman also reported the common reactions from those with MbI once they have been exposed offering some insight into the disorder. The reactions tend to be very disparate:

- Protesting their innocence.
- Blaming other users i.e. if they had been more supportive they would not have had to make up stories.
- Leave the group and carry on their deception elsewhere.
- Admit that they were lying but cannot explain their behaviour.
- Admit that they were lying and mock other users for believing them.
- Cyber stalk and threaten those who have exposed them.

Another impact of MbI is more obscure and relates to internet mediated research as was highlighted by Lawlor and Kirakowski (2014). Using the internet for data collection is becoming increasingly popular as it offers many advantages including access to large and diverse populations, enabling collection of sensitive data, low cost, efficient, 24/7 access and elimination of experimenter effects (Russell & Purcell, 2009). Its popularity is evident by the number of online communities that include sections specifically for researchers to recruit participants and the number of textbooks for researchers which now contain specific sections or whole books dedicated to internet mediated research methods (Fielding, Lee & Blank, 2008; Bryman, 2012; Somekh & Lewin, 2011; Hewson, 2003). There are also intense ethical debates around informed consent and public/private spheres on the internet as further evidence of the solidification of internet mediated research (McKee & Porter, 2009; Buchanan, 2004; Whiteman, 2012). However, internet mediated research is not without its

disadvantages and one of the main disadvantages is the anonymity of participants compromises data (Fiske, Gilbert & Lindzey, 2010). Despite this risk, the inclusion of MbI sufferers in data has received little attention especially for those conducting health related research.

1.3.6 DETECTING AND MANAGING MBI

Research has been conducted into the features associated with MBI by Feldman (2000). Feldman (2000) compiled a series of cues that are associated with MBI and notes that detection of MBI is difficult as in many instances there is a mix of facts and lies used in the ruse:

- Consistently copies from other posts, text books or health related websites.
- The posting behaviour (length, frequency, duration) is inconsistent with the severity of illness being claimed e.g. having a heart attack and posting from the back of an ambulance.
- The characteristics of the illness and treatment are caricature like due to misconceptions.
- Oscillating from near to death to miraculous recovery.
- Fantastical personal claims which are contradicted or disproved.
- Constant dramatic events which escalate when attention towards them wanes.
- Complains that the other members are inattentive and this is compromising their health.
- Avoiding telephone contact by offering odd excuses.
- Feigned blitheness about a crisis which will instantly attract attention.
- The use of 'sockpuppets' i.e. people posting on their behalf who happen to have the same unique writing style.

The problem with these cues is that they were derived from a handful of case studies and some are open to interpretation. Without solid evidence, there would be naturally a reluctance to accuse someone of faking an illness, if wrong a vulnerable person is being falsely accused of a heinous act. This is further compounded by our naturally leaning towards the truth bias, we expect the truth. Even people who know what cues to look out for or are even formally trained in the cues of deception, they are still

notoriously poor at detecting deception (Burgoon, Guerreo & Floyd, 2010).

If a case of MbI is exposed, the question remains as to what strategy is best for dealing with such people. One suggestion has been to publicly expose the identity of MbI sufferers. Feldman and Peychers (2007) reported on a case where a large online self-help website was abused by a member claiming to have experienced multiple life-threatening issues. Suspicions were raised due to the contradictory information he shared and his abrasive and threatening attitude once confronted. In response, several members set up a website with the aim of exposing his behaviour and identity to the public. The faker took legal action to try and close the website based on defamation. The defence was that the website was in the public interest and was completely truthful. The judge asked for the plaintiff's medical records to be allowed which had previously been sealed. On this request, the plaintiff withdrew and the case was dismissed with prejudice. Therefore, members of online communities have the power to expose deceptive behaviour and the identities of perpetrators without fear of retribution, provided the claims are correct. In addition, Pulman and Taylor (2012) suggest that MBI should be considered as a cybercrime rather than an accepted hazard of participating in online community that cannot be controlled. Particularly in cases where incorrect medical information is disseminated as part of the ruse which subsequently results in the worsening of a health condition or fatality. In such instances perpetrators should be pursued through their IP addresses. In support, Feldman and Peychers (2007) argue that pursuing lawsuits is the ultimate way of turning the tables on those with FD, the same could be applied to MBI. Another novel solution to stem the medical misinformation associated with MBI was suggested by Witney, Hendricks and Cope (2015). They argue that online support groups should employ a health professional who would can provide evidence based health information and correcting any misinformation posted by members.

1.4 DIGITAL DECEPTION

1.4.1 ADVANTAGES OF THE ONLINE ENVIRONMENT

MbI is not just situated within FD but also must be understood within the wider realm of digital deception where these ruses are played out. The internet provides the ideal environment for FD sufferers because it is low risk and offers a high chance of success.

The different features relevant to deception were categorised by Kaplan and Haenlin (2010) when they developed a method of classifying different social media services based on social presence (acoustic, visual and physical immediacy of the communication medium), media richness (the amount of information that can be conveyed to reduce ambiguity and uncertainty), self-presentation (control how others perceive them by creating an image) and self-disclosure (conscious or unconscious revealing of personal information). Tsikerdekis and Zeadally (2014) used this model to assess different social media services and the level of difficulty involved in deception as well as the success rate. The social media services used by MbI sufferers (blogs, social networks, and content communities) are regarded as being low on risk and high on success. As they allow for the manipulation of identity easily (low self-presentation/self-disclosure) and increased control of the content because it is text based and asynchronous (low social presence/media richness). Thus, the online environment is ideally suited to MbI sufferers.

The MbI perpetrator's ability to successfully deceive is even further facilitated online by what Donath (1999) refers to as conventional and assessment signals. Donath (1999) applied the "Handicap Principle" proposed by biologist Zahavi (1977) to online identity deception. The principle states that reliable signals are evident by the congruence between the cost of the signal and the traits displayed by the signaller. Assessment signals follow the handicap principle and so are reliable, an example of an assessment signal would be a person signalling they are wealthy by driving a Ferrari, to have the Ferrari the person must possess the trait of wealth therefore the signal is reliable. Conventional signals do not follow the handicap principle. The signaller does not necessarily have to possess the trait associated with the signal. A person could signal they possess the trait of wealth by having a Ferrari key ring, while this signal is associated with wealth it is clearly not full-proof, most anyone can buy a key ring without much cost. The problem with online communication is that assessment signals are very limited due the medium and therefore people must depend on the less reliable conventional signals which can be produced by those displaying genuine identities and those involved in online identity deception. Hancock (2011) describes how these signals are translated online. Online conventional signals are the information included in a profile or message that is taken at face value e.g. I'm wealthy and my name is 'Richie-rich'. Online assessment signals are less available online, they

reveal a person's real identity e.g. phone number, professional profile, or knowledge only someone with their expertise would have. Donath (1999) found that because online communication largely depends on conventional signals which cost little to attain compared to assessment signals, a user could easily afford to have multiple personas. The same advantage of online communication is also exploited by MbI sufferers.

1.4.2 TYPES OF ONLINE IDENTITY DECEPTION

Hancock (2011) defines digital deception as the 'intentional control of information in a technologically based mediated message to create a false belief in the receiver of the message'. Digital deception is divided into two categories, identity-based digital deception, and message-based digital deception. Identity deception involves misrepresenting a person or organisations identity. More specifically when related to people, it is pretending to be a different person or kind of person than one is (Donath, 1999). Message deception is transmitting false information in online communications. Identity deception however, is not independent of message deception as the communications transmitted by those misrepresenting an identity will inherently contain false information. Identity deception can be further divided into three groups, category deception, impersonation, and identity concealment (Donath, 1999).

Category deception involves the provision of false demographic details such as sex, age, employment, personality etc. (Donath, 1999). Category deception has become commonly known as 'catfishing' in popular culture. The origin of the term derives from a documentary called 'Catfish' (Joost & Schulman, 2010) that exposed a case of online category deception. One contributor, who was married to the lady who created the fake online persona around which the documentary was based, described how catfish are put into tanks with cod when they were transported between Alaska and China, this is because they chase the cod keeping them fit and thus ensuring the quality of the meat. Her husband equates the function of catfish with that of those involved in online category deception, believing their deceptions make online interactions more interesting and vibrant (Palmer, 2012). However, many would argue against this positive view of category deception.

Now, regarding impersonation, Tsikerdekis and Zeadally (2014) describe it as occupying an already existing identity (rather than creating one) to obtain information from their peers. The target is not necessarily the person whose identity has been stolen but rather the peers in their social network. They gain their trust and then exploit it for information. As there is no reliable way of verifying the identity of members of online communities it is relatively easy to do. However, Tsikerdekis and Zeadally (2014) argue that long-term success is unlikely because the real owner of the identity will eventually become aware of the ruse or others will become suspicious, as the impersonator will not be able to maintain congruency between behavioural cues and the impersonated identity. Concealment is different to both category deception and impersonation because it involves withholding information rather than providing information which is misleading (Whitty & Joinson, 2009). MbI sufferers appear to be involved in both category deception and impersonation, concealment is less applicable. They have been known to create whole new identities or steal photos from other people profiles and pass them off as their own, particularly those of sick children (Monroe, 2016).

1.4.3 DECEPTIVE TACTICS

Those involved in online deception are believed to employ several deception tactics to create a false cognitive representation of what is occurring. Unfortunately, there is no direct research that examines the tactics used by those involved in online identity deception, however Grazioli and Jarvenpaa (2003) did examine strategies that are available to businesses and consumers, these could also be utilised by individuals such as MbI sufferers. Grazioli and Jarvenpaa (2003) describe the deceptive interaction as a deceiver manipulating the victim's environment, to foster a false representation of the situation, so that the victim is manipulated into behaving in a way they would be unlikely to if they had not been manipulated. They manipulate the victim's environment to produce these desired outcomes using a variety of tactics. Grazioli and Jarvenpaa (2003) believe that those most relevant to online deception include, masking (holding back or eliminating information), dazzling (misrepresenting information), decoying (distracting from what is really happening), mimicking (impersonating or assuming another's identity), inventing (making up information), relabelling (using misleading information), and double play (blaming the victim for taking advantage of

them). These tactics can be divided into two groups, those that obviate the creation of a correct understanding of the deception core (masking, dazzling, decoying) and those that create a false understanding of the deception core (mimicking, inventing, relabelling, double play). It's plausible that all these deception tactics could be used by MbI sufferers, however given that they create fictional stories, those most relevant are likely to be tactics which create a false understanding rather than obviate.

1.4.4 IDENTIFYING ONLINE IDENTITY DECEPTION

The multiplicity of deception techniques available and the lack of deception cues makes identifying online identity deception challenging. However, online users have found novel ways to use whatever cues may be available to them and acting as their own detectives to validate a person's identity. For example, a study by Ellison, Heino and Gibbs (2006) found that users were adapting to what cues were available to them and how they could utilise the available cue in strategic ways to assess the credibility of online identities as well as adding credibility to their own. These two goals were often reciprocal, as by thinking about what cues they could use to increase the credibility of their own identity, these same cues could be used as indicators of online identities. For example, one woman decided to choose a photo of her standing up so that there was no doubt about her physique. As she believed that those who were heavy might choose a more ambiguous photo to hide their weight. Therefore, when viewing online profiles if she wanted to increase her certainty that a potential partner had a slim physique she would only contact those with photos where the person was standing as she had done to demonstrate her physique. A similar pro-active strategy is being used by members of online communities when they become suspicious that there is an MbI sufferer within their group. An online group led by Taryn Wright investigate suspected cases of MbI and has uncovered dozens of cases by learning to use what assessment cues are available to them and conducting extensive research to unearth them (Monroe, 2016). For example, cues may include discrepancies in pictures e.g. claiming to be bald due to cancer treatment but still having eyebrows or claiming to be on steroids but not having the associated swollen face.

A new emerging method of identifying online identity deception is Natural Language Processing (NLP), and specifically text classification using supervised machine

learning techniques. Bird, Klein and Looper (2009) describe text classification as the use of machine learning techniques (decision trees, Naïve Bayes and maximum entropy) to identify features within a text (word structure, word frequency) that correlate with the category of the text that needs to be classified (topic, gender, authorship, sentiment etc.). Once distinguishing features associated with the target category have been identified, the classifier can automatically identify what category any new text it encounters belongs to. Text classification is ideally suited to detecting online identity deception as the primary mode of communication is text based and there is a large corpus available to train and test a classifier. Furthermore, a text classifier also offers advantages over human judges who do not perform better than chance whether they are trained or untrained in deception cues (Bond & DePaulo, 2006). For example, a study by Ott et al. (2011) found that text classifiers outperformed humans in detecting fake online reviews with an accuracy rate of 90% for the classifier compared to 62% for human judges. There are two theories as to why people are bad at detecting deception. The first is truth bias, the tendency to believe that a message is true regardless of the veracity of the message (Levine & Kim, 2010). The second is that the cues may not be obvious enough to be perceived and accurately interpreted by humans (DePaulo et al., 2003; Hartwig & Bond, 2011). This is further confounded when there are limited cues available, such as online. A text classifier therefore is well suited to the online environment and offers advantages over humans, it is unbiased and can notice subtle linguistic cues which may go unnoticed by the human eye. These advantages have led to text classifiers being successfully used in the online environment to detect numerous forms of online manipulations (fake reviews, harassment, and cyberterrorism) and identity deception (sexual predators, sockpuppets and deceptive dating profiles) (Afroz, Brennan & Greenstadt, 2011; Ott, Cardie & Hancock, 2012; Ott et al., 2011; Alowibdi et al., 2013).

1.5 SOCIAL LANGUAGE PROCESSING (SLP)

Text classification provides the technical blueprint from a computational point of view. However, it completely ignores the theoretical underpinning of classifiers, particularly in relation to choosing features and interpreting the relevancy of discriminating features. This is where Social Language Processing (SLP) has bridged the gap. SLP is a research framework developed by Hancock et al. (2010) for

conducting text classification that brings together social and psychological theory with computer science. SLP has its own three stage recursive process for developing a classifier that includes linguistic feature identification, linguistic feature extraction and classifier development. Both linguistic feature extraction and classifier development cover the computational techniques involved in developing a classifier. The innovative stage in SLP, that draws in social and psychological theory, is linguistic feature identification.

SLP posits that the selection of potentially discriminating features should be based on social and psychology theories. Essentially SLP argues that any text classification task should begin with a hypothesis as to what features will be discriminatory and why. This is in stark contrast to a purely computational approach that does not begin with a hypothesis but rather delves straight in by including random features that are not necessarily grounded in theory. Using deception as an example, Hancock et al. (2010) argues, that based on the literature, features such as first person singular pronouns (reflect the deceivers attempt to deflect attention from themselves when lying), matching linguistic styles between the deceiver and their target (deceivers emulates their targets linguistic style to try and gain approval) and the amount of cohesion in a deceptive text (deceiver may display above average cohesion as their story has been rehearsed or they may display below average cohesion as they struggle to keep track of their lies) should be included given their theoretical backing.

However, SLP also allows for innovative findings by leaving room for the unearthing of additional linguistic features beyond those specified in the hypotheses. Hancock et al. (2010) regards this space for discovery as particularly important in cases where there is no theoretical indication as to what linguistic features may be associated with the target class. In this way SLP can help further existing knowledge and inspire future studies by discovering new linguistic features, and based on the literature, retrospectively theorising how they may relate to the social and psychological constructs of the individuals in the target class. SLP therefore does not just place value of the classifier itself, but also on what the discriminatory linguistic features used by the classifier tell us about the social and psychological constructs of the individuals in the target class.

Another facet of SLP is the importance placed on acknowledging a key limitation of text classification, that is context. Context is often overlooked by purely computational approaches where accuracy is the only goal. Hancock et al. (2010) argues that it is essential to acknowledge what domains and contexts the classifier can be regarded as reliable in. This is because the data that was used to train the classifier may not be equivalent to the real-world data that the classifier is being applied to. However, by understanding what the discriminatory linguistic features say about the social and psychological constructs of the individuals in the target class, it is possible to gauge the domains and contexts that the classifier can be reliably used to. For example, if a classifier is trained to identify the gender of a writer by training it on text written by males and females, it could not be reliably used to detect writers who try to obfuscate their gender. This is because gender deception will have its own distinct set of linguistic features that reflect the deceivers own social and psychological constructs rather than the discriminatory linguistic features simply matching those associated with text written by males and females.

1.5.1 TEXT CLASSIFICATION APPLIED TO ONLINE DECEPTION

The four stages (pre-processing, feature extraction, classifier training and classifier testing) provide a standardised method for developing a text classifier and have been applied to detect numerous forms of online manipulation and identity deception including opinion spam, online harassment, sockpuppetry and sexual predation which are reviewed in the following sections. Although these studies don't explicitly use the SLP research framework, they are in the same vein as they do attempt to integrate psychological and social theory when choosing and interpreting the relevancy of discriminatory features.

1.5.1.1 OPINION SPAM

Given the economic cost to businesses, one of the most researched uses of text classification is calculating the veracity of online reviews, also referred to as opinion spam, by categorising reviews as either authentic or manipulative. Across the studies there does not appear to be a consistent set of features that is associated with opinion spam. However, in comparison to low level content based features, high level context based features consistently improve the accuracy of the text classifiers. Karami and

Zhou (2015) used content features including the presence/frequency of unigrams and bigrams to train classifiers with three different machine learning techniques, logistic regression, support vector machines and Naïve Bayes. The detection accuracy across the classifiers was no greater than chance at 50%. Character n-grams (continuous sequence of n items from a given sequence of text), although similarly content features, do produce higher detection accuracy rates of between 76% and 90% using Naïve Bayes and SVM, compared to word n-grams (Hernandez Fusilier et al., 2015; Cagnina & Rosso, 2015). Kanaris et al. (2006) explains that although both word and character n-grams are content based, character n-grams have a distinct latent advantage by tapping into context through content, such as lexical, word class and punctuation usage, all of which are associated with authorship. Character n-grams, can therefore be used to distinguish authentic and manipulative reviews based on writing style, as the writing style of a deceiver is hypothesised to be different from that of an authentic reviewer. Using more direct context based features, such as the psycholinguistic dimensions in the LIWC, in combination with content features is arguably the best method of detecting opinion spam. Ott et al. (2011), in a large-scale study of feature sets, reported that the best performing classifier at 90% accuracy was modeled using a combination of both context and content features.

1.5.1.2 ONLINE HARASSMENT

Text classification has also been used to detect online harassment but has required a more innovative approach given the uniqueness and complexity of the behavior. The baseline approach is the identification of features that are likely to be indicative of harassment. The seminal work by Yin et al. (2009) examined three groups of features:

Local features – specifically, Term Frequency/Inverse Document Frequency (TFIDF), the importance of a word in a single post in relation to its frequency of use within the larger corpus.

Sentiment features – based on observations, increased usage of foul language, second person pronouns and all other pronouns.

Contextual features – sentiment and local features in isolation are not always indicative of harassment, therefore posts are compared to other posts around them as

often a harassment post will stand out or as other users respond a cluster of harassing posts will co-occur.

Individually these three features produce low levels of prediction accuracy, however when combined a prediction accuracy of between 31% and 44% is achieved, depending on the dataset used. Compared to the detection of opinion spam this result is relatively poor. Subsequent studies have tried to improve the accuracy by using different approaches. Dadvar et al. (2012) used a gender based approach, as harassment language varies per the gender of the harasser. Using the combination of features identified by Yin et al. (2009), two classifiers were trained separately using SVM on male and female datasets. Using a gender specific classifier improved the accuracy of detection by 15% compared to the baseline. Dinakar, Reichart and Lieberman (2011) used a label specific method by developing three topic sensitive classifiers to identify conversations around race, sexuality, and intelligence. These three topics are of a personal a nature and are often utilized by harassers to cause the greatest offense. The topic sensitive classifier, in combination with profanity and negative connotation features, can predict online harassment with an accuracy of between 70% and 80% compared to 63% when the data is merged. Dadvar et al. (2013) moved beyond the detached content of messages (profanities, number of emoticons, personal pronouns) and cyberbullying features (sensitive topics, cyberbullying words) to include user based features attached to these messages, such as the history of the user's activities (pattern of offensive language use, length of comments, use of profanities) and the age of the user. Including user based features in combination with the content and cyberbullying features resulted in a 4% increase in accuracy, from 60% to 64%.

1.5.1.3 SOCKPUPPETS

One form of online identity deception that has been successfully detected using text classification is sockpuppetry. Weckerle (2013) broadly describes sockpuppets as people who create false identities online to intentionally deceive or manipulate others. Wikipedia has a high prevalence of sockpuppets, which is problematic as they often deface existing articles and/or create bogus articles of their own despite repeated bans. Solorio, Hasan and Mizan (2013a) trained a text classifier to detect sockpuppets on Wikipedia with an accuracy of 69% by using authorship attribution including stylistic,

grammatical, and formatting features that discern individual authors. The accuracy was increased to 72% when two timing features were added, the time and the day of the week edits took place. In a second study by Solorio, Hasan and Mizan (2013b) a larger corpus of 623 cases was used, compared to the 77 cases used in the previous study. A different framework was also used, instead of training the classifier with individual comments as instances, a pair of editors is considered an instance and the classifier must determine whether the pair is a match or by two different editors. Using the same authorship attribution and timing features an accuracy of 73% was achieved, showing that the classifier maintained its robustness even on a larger corpus. The inclusion of the non-verbal behavior of users as features does appear to improve the accuracy of classifiers and is particularly applicable for detecting sockpuppets on Wikipedia given the availability of activity log files (Kuruvilla & Varghese, 2015). Tsikerdekis and Zeadally (2014) argue that classifiers trained using non-verbal features are superior to those trained using verbal features. This is mainly because they are computational more efficient and cost effective as they require less data to be processed, which is important when dealing with high traffic websites. Additionally, deceivers may be less able to manipulate non-verbal features, be less aware of their activity log being monitored and even if they are aware, be less able to alter their activity to evade detection. Tsikerdekis and Zeadally (2014) examined all the non-verbal features available from the activity log files of Wikipedia users and found that the most important variables in their classifier to predict sockpuppets, which achieved a prediction accuracy of 71%, were the distribution of edits, higher probability of contributing to specific pages, more active and specifically more active in pages other than the article's page and longer duration between postings and revisions. Yamak, Saunier and Vercouter (2016) also examined the use of non-verbal features to detect sockpuppets in Wikipedia and developed a classifier, using the same Random Forrest algorithm as Tsikerdekis and Zeadally (2014), with an accuracy of 99.5%, adding further support to the superiority of non-verbal features. The significantly higher accuracy rate achieved by Yamak et al. (2016) is attributable to two new features that were included, the frequency of reverts made by other users and the average contribution per article. Kumar et al. (2017) extended the use of non-verbal features beyond activity to include a new community feature, in addition to the regular activity and linguistic features. That is the interactions between the user and the community. As it was hypothesized that sockpuppets would be treated more severely than other

users. Kumar et al. (2017) tried to detect sockpuppets in online discussion communities by investigating if it was possible to discern a sockpuppet from a regular user and if pairs of sockpuppets could be identified. The community features used to train the classifier were the fraction of downvotes a user received on their posts, the fraction of posts that were reported and deleted and if a user had been blocked. The classifier for detecting sockpuppets achieved 68% accuracy, however the addition of community features had no impact once removed from the classifier. An accuracy rate of 91% was achieved for detecting sockpuppet pairs, but once again the community features performed the poorest with an individual prediction accuracy of 56%.

1.5.1.4 SEXUAL PREDATORS

A second form of online identity deception that has been detected using text classifiers is sexual predation. Pender (2007) used low level features (unigrams, bigrams, and trigrams) to try and detect sexual predators who believed they were conversing online with underage victims, but were in fact pseudo-victims i.e. a volunteer posing as a child. The text classifier with the highest accuracy of 94% was trained using trigrams. On closer examination Pendar (2007) found that the trigrams were highly correlated with the 'sliminess' of the predator, as rated by the pseudo-victims. It is hypothesized that the predictive power of trigrams underscores the three-stage communication strategy used by sexual predators (building trust, assessing if this is a possible victim, introducing sexually explicit language/images) and not linguistic features that set sexual predators apart. Escalante (2013) investigated the three-stage communication theory by building a chain classifier to detect each of the stages, on the basis that the language used in each stage is different, and then linked the classifier to the complete text. Interestingly, the classifier trained using only stage one (gaining access) outperformed the global classifier that used all three stages by 12%. Escalante (2013) argues that this provides evidence that how a well-intentioned person approaches a child differs from a sexual predator. The predictive importance of the initial approach by a sexual predator has also been supported by Gupta, Kumaraguru and Sureka (2012) who found that the initial friendship forming stage accounts for 40% of the lines in a pedophile conversation. Both Peersman, Daelemans and Van Vaerenbergh (2012) and van de Loo, De Pauw and Daelemans (2016) used a different approach by predicting age and gender based on the premise that pedophiles provide false age and/or gender information to get closer to children so they can groom them. Both

studies developed classifiers with high accuracy using data obtained from a social network and training the classifier with word and character n-grams. However, a significant problem is whether these classifiers could be used in a real world setting where a pedophile will intentionally try to obfuscate their age and gender. Bogdanova, Rosso and Solorio (2012) used the psychological characteristics of pedophiles as their guide for choosing salient features to train their classifier. These characteristics include feelings of inferiority, isolation, loneliness, low self-esteem, and emotional immaturity. These high-level features were then compared with low level features i.e. word and character n-grams. The results showed that using high level features based on the psychological characteristics of pedophiles produced a classifier with a very high accuracy rate of 94% compared to a classifier using low level features which performed close to chance. One exception was the classifier trained using tri-grams which had an accuracy of 72%. Pender (2007) also reported on the predictive value of tri-grams as being indicative of the communication strategy used by pedophiles.

Although the above examples of text classification applied to online deception did not explicitly use SLP, they do illustrate the value of exploring the relationship between discriminate features and the psychological and social constructs underpinning the individuals within the target class, which is at the core of SLP.

1.6 RESEARCH AIMS

The aim of the thesis is to develop a method to detect MbI. Such a method could be used by internet mediated researchers to help ensure that they do not unwittingly collect data from MbI perpetrators which could lead to spurious findings. A method to detect MbI could also be used by moderators, given the emotional distress MbI causes to those who are duped and the disruption it causes to online communities. The distress and disruption caused is particularly acute for online support groups where members are likely to be emotionally vulnerable. Presently assessment cues are being used to detect MbI, however they are susceptible to human judgement bias and take time to unearth. To develop an alternative method of detecting MbI an in depth understanding of the disorder is required. This knowledge is a prerequisite for firstly choosing what method of detection is most appropriate and secondly for selecting and interpreting the cues that may be indicative of the deceptive behaviour. For example,

in the case of using the SLP research framework for developing a text classifier, knowledge of the psychological and social constructs underpinning the individuals within the target class is essential when it comes to selecting and interpreting the discriminatory linguistic features used by the classifier.

Therefore an additional aim of the thesis is learn more about the aetiology of MbI and its offline version FD. This is because at present the baseline knowledge of these disorders is scant. What is known about FD disorders is purely speculative and based on a handful of observational case studies. The value of the findings from these case studies is also limited as the basic questions around diagnosis, aetiology and treatment have yet to be adequately answered. This has caused issues for clinicians who have little evidence based information about FD at their disposal and what is known does not appear to facilitate the diagnosis and treatment of FD. The complete reliance on observational case studies has arguably led to this limited understanding of FD. Furthermore, the findings of these case studies is questionable given their poor practical application value for clinicians. This incongruence indicates that there may be a mismatch between our limited theoretical understanding of FD based and the reality of FD for sufferers. Therefore, the implications of learning more about FD and MbI goes beyond the development of a classifier to detect MbI, but will also contribute to our understanding of the disorders.

Given the litany of issues, the question arises as to why more first-hand information has not been collected about FD and MbI. The main impediment has been access to FD sufferers and MbI perpetrators. FD sufferers and MbI perpetrators are difficult to recruit as participants because they are likely to be afraid of coming forward and having their deceptive behaviour exposed to those around them. Exposure would leave them vulnerable to the negative stigma surrounding the false occupation of the sick role. This is stigma is due to FD sufferers completely violating the social contract of the sick role leaving people who are duped feeling angry and frustrated. Furthermore, once a person is exposed as having FD, they will inevitably lose their coping mechanism as it will be impossible to occupy the sick role with any level of credibility. Therefore, because of the potential disadvantages of exposure by participating in studies, FD sufferers do not tend to engage with researchers. This has led to the reliance on case studies which has generated speculative information about FD with

little practical application value.

The aim therefore is to build up a more comprehensive baseline knowledge of FD and MbI that can be used to inform the development of a method to detect MbI and deepen our understanding of the disorders. This will be achieved by moving beyond observational cases studies that are largely speculative and instead using first-hand accounts of FD and MbI. These accounts will be collected from those who suffered from FD and those who have encountered MbI online. Given the difficulty in recruiting people to participate in studies of FD and MbI, the pre-existing posts in online communities will be used as the primary source of data. These posts will be collected from an online community for FD sufferers and from a variety of online communities that have experienced MbI. The data obtained will then be analysed using grounded theory and the findings will be used to either corroborate existing theories or develop new theories. The knowledge garnered from these first-hand accounts will be used to develop a method to detect MbI. By providing the information necessary to choose the most appropriate method of detecting MbI and providing direction as to what cues may be indicative of MbI and how discerning cues can be interpreted in terms of the psychological and social constructs of MbI perpetrators. In addition, the findings from these studies will further our understanding of FD and MbI beyond what has been found through observational case studies and will hopefully inspire further research in this area.

In conclusion, the aim of the project is to develop a method to detect MbI and in doing so develop a knowledge base of both FD and MbI that is grounded in first-hand accounts, a perspective that is currently lacking in the literature. The creation of this knowledge base will provide direction as to what method of detecting MbI is most appropriate and will give theoretical context to the salient cues associated with MbI. The salient cues associated with MbI may also give rise to new insights into the disorder. It is hoped that this knowledge base will lead to a new understanding of MbI and FD that will benefit both clinicians and sufferers. To achieve these aims three studies were conducted, the aims of which are described below:

Study One – *When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder*. The aim of study one is to try and fill a gap in the current

FD literature which has been left by the lack of first-hand accounts from FD sufferers themselves. This gap is inhabited by several questions including; whether FD sufferers experience any negative symptoms; whether the motivation for occupying the sick role is conscious or unconscious; whether sufferers are perturbed by their own behaviour; and whether sufferers want to recover from their disorder and if so why do they not seek professional help.

Study Two – *Claiming someone else's pain: A grounded theory analysis of online community user's experiences of Munchausen by Internet.* The aim of study two is to similarly move beyond observational case studies, but instead shifting the focus closer to MbI. First-hand accounts of those who have interacted with MbI perpetrators in online communities will be used to elucidate the presentation of MbI in online communities and the associated behaviours.

Study Three – *Detecting Munchausen by Internet: Development of a classifier through machine learning.* The final study will integrate the findings from studies one and two to aid in the development of a method to detect MbI. The findings from studies one and two will inform what method will be used and aid in choosing and interpreting the discriminatory cues associated with MbI by situating them within the wider context of the social and psychological constructs associated with MbI. The cues taken on their own may also contribute to our understanding of MbI.

The following section includes information about the qualitative research methodology used in studies one and two that were beyond the scope of the published papers. While the remaining chapters of the theses present the three papers that were obtained from the above studies and a concluding chapter outlining both the practical application value and theoretical implication of the three studies including the corresponding limitations and areas for future research.

1.7 QUALITATIVE RESEARCH METHODOLOGY

This section describes the research methodology Grounded Theory (GT) that was used in both study one, *When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder,* and study two, *Claiming someone else's pain:*

A grounded theory analysis of online community user's experiences of Munchausen by Internet. An overview of GT and the different schools within the paradigm is provided. The suitability of the GT methodology for the two studies is discussed and decision to use the Constructivist Grounded Theory (CGT) methodology is explained. The methodology of CGT is also described to provide the in-depth details of how the framework was applied to the data.

1.7.1 GROUNDED THEORY OVERVIEW

Grounded Theory (GT) was developed by Glaser and Strauss (1967) and grew from the need to reconcile two conflicting methodological necessities (Kelle, 2005). That is, can purely new theories be generated without the need for drawing on pre-existing theories, a requirement of empirical research. Glaser and Strauss wanted to develop a method that would transcend the hypothetico-deductive method where a hypothesis or theory is required prior to data collection. Within Sociology, the dominant form of research was theory verification. However, Glaser and Strauss advocated exploring new theories and concepts that were tailored for the area of research under investigation. This led to the development of GT, "a general methodology of analysis linked with data collection that uses a systematically applied set of methods to generate an inductive theory about a substantive area" (Glaser, 1992, p. 16). According to Glaser and Strauss (1968, p.3), these theories should (1) enable prediction and explanation of behaviour; (2) be useful in theoretical advances; (3) be applicable in practice; (4) provide a perspective on behaviour; (5) guide and provide a style for research on specific areas of behaviour; and (6) provide clear enough categories and hypotheses so that crucial ones can be verified in present and future research. Evans (2013) argues that while Glaser's (1992) definition of GT and the scope of these theories described by Glaser and Straus (1968) are largely uncontested amongst researchers, there are differences in opinion as to how data should be handled and analysed. These differences in opinion led to a split within GT resulting in two sets of schools, Glaserian vs. Straussian and Objectivist vs. Constructivist (Charmaz, 2008).

1.7.2 BASIC TENETS OF GROUNDED THEORY

Despite the different schools within GT, there are underlying concepts that all schools converge on. Kenny and Fourie (2015) describe these five concepts as theoretical

sampling, saturation, comparative analysis, memos, and substantive versus formal theory which are elucidated by Glaser and Strauss (1967). Firstly, theoretical sampling refers to decisions regarding data collection not being pre-planned, but rather guided by the research process that may reveal the need for more data. The initial coding and categorising of data may uncover gaps that require further data collection to fill. The simultaneous collection, coding and analysis of data may change the direction and scope of the research resulting in the need for new sources of data to pursue these unanticipated avenues. As the hypothesis starts to take shape, further data may need to be collected to add extra credence. Secondly, saturation is reaching a point whereby collecting more data will not lead to any new theoretical categories or properties associated with these categories. Essentially there are no more theoretical insights to be gained by collecting new data as the same information is cropping up repeatedly. Thirdly, comparative analysis requires the researcher to be continuously involved in constant comparison. This iterative process means comparing codes with codes, codes with categories, categories with categories and theory with the literature. Through constant comparison the researcher can develop a theory that is closely aligned with the data and is homogenous and plausible. Fourthly, memos are notes that provide the researcher with the opportunity to reflect and record the deliberations that occurred throughout the research process. The memos also facilitate theorising as they serve to exemplify the categories and provide an overview of the relationships between the categories. The memos can also facilitate the writing up process by reminding the researcher of the network of decisions that were made to enable them to reach a conclusion. Fifthly, a strong distinction between substantive and formal theory is made whereby GT researchers should concern themselves with developing only one type of theory to ensure its accuracy. Substantive theory is applicable to the specific research area under investigation and requires further study to test its applicability in wider settings, whereas formal theory is a more general and can be applied to large range of settings. These five concepts provide common ground for the different schools within GT however there are substantive difference between the schools that require a researcher to position themselves within the schools according to their own philosophical positions, primarily choosing between Glaserian, Straussian and constructivist GT.

1.7.3 GLASERIAN VS. STRAUSSIAN

Cooney (2010) describes the origins of the split between Glaser and Strauss as being due to conflicting opinions about data analysis. The process of data analysis had been sketched out in Glaser and Straus's (1967) first publication that introduced GT to researchers. In their subsequent endeavours to flesh out the processes involved in data analysis their differences in opinion became apparent. Glaser adhered to the original version of GT while Straus wanted to reformulate the original version. Glaser and Straus went their separate ways in 1992 after Strauss and Corbin's (1990) book titled *'Basics of qualitative research: Grounded theory procedures and techniques'* served to highlight the contrasts in their respective approaches to GT. According to Gasson (2009), there were two issues of contention, ontological and methodological. The ontological issues were whether theory generated from GT is independent of the researcher and limited to the circumstances of the research area under inquiry, or whether theory is socially situated and subjective. Glaser supports the former and this is referred to as critical realism, whereas Strauss supports the latter which is referred to as relativism. Glaser contends that the theory is embedded within the data and can be discovered, whereas for Strauss and Corbin the theory is interpreted based on the consensus of the actors involved. Glaser also stresses the need for triangulation to verify a theory developed from GT, however for Strauss constant comparison of data samples within the GT process is enough to verify a theory. Breckenridge et al. (2012) argues, that despite the rift between Glaser and Strauss with regards to their respective philosophical positions, Glaser (2002) argued that it was the method that mattered and that GT was ontologically and epistemologically neutral. This means that GT is a versatile general method which can be fitted to a variety of epistemological and ontological stances as directed by the researchers own philosophical position and that which is most appropriate for the data and research area under investigation. However, the ontological and epistemological neutrality of GT has been widely criticised because within qualitative research stating a philosophical position is intrinsically linked to how the research method will interpret the data (Holton, 2010).

Despite the lack of consensus over philosophical positions, the methodological differences between Glaser and Strauss are clearer. According to Kenny and Fourie (2015) the differences begin with use of literature. Due to Glaser's critical realist

stance that there is an objective reality, the onus is on the researcher to remove themselves from the research. This means negating all pre-existing academic and personal knowledge of the research area under investigation. Therefore, consulting with the literature is discouraged as it would inhibit the theory from emerging naturally from the data free of preconceptions and biases. In contrast, Strauss's relativist stance argues that the researcher cannot be separated from the research. Therefore, literature should be consulted during the research process but not to the extent that it stifled the creativity of the researcher in identifying novel theories. Strauss and Corbin (1990) advocated an *open mind* over an *empty head*. Gasson (2009) describes the second methodological issue as hinging on how data analysis within GT should be structured to create a systematic method. Strauss advocated a highly-structured approach by introducing an additional step in the coding process, *axial coding.* According to Böhm (2004), the addition of *axial coding* helps to clarify and discriminate concepts so they can be transformed into categories. Theory development then occurs by examining the relationships between the axial categories and their related concepts. Strauss and Corbin developed a *coding paradigm* that researchers can use to establish what the relationships are thus providing a skeleton for the development of theories. Broadly speaking, the *coding paradigm* includes temporal and spatial, cause and effect, means and ends relationships all situated within argumentative and motivational connections. The *coding paradigm* essentially provides pre-defined options for the researcher to choose amongst when asking questions including under what conditions did the category arise and under what context, what action/interaction strategies were used to carry out the processes and the consequences of using these strategies (Kendall, 1999). Glaser (1992) argued that using these predefined coding paradigms would result in categories being 'forced' on the data instead of allowing them to 'emerge'. Gasson (2009) succinctly sums up the methodological difference between Glaser and Strauss. Glaser supported dynamic theories of human agency and interaction, whereas Strauss and Corbin advocated a more static approach using a set of predefined paradigms to frame the theories.

1.7.4 OBJECTIVIST VS. CONSTRUCTIVIST

Constructivist Grounded Theory (CGT) essentially puts aside the philosophical positions of both Glaserian and Straussian GT, jointly referred to as Objectivist

Grounded Theory (OGT), but retains their flexible strategies for analysis (Charmaz, 2008). CGT was brought about by Charmaz (1995) who wanted to depart from the positivist position associated with Glaserian and Straussian GT. Charmaz (2003) describes how CGT "takes a middle ground between postmodernism and positivism, and offers accessible methods for taking qualitative research into the 21st century" (p. 250). However, it has been argued that Charmaz's CGT is very closely aligned with postmodernist relativist ontology rather than being in the middle ground (Kenny & Fourie, 2015). Charmaz was critical of the positivist assumptions that there is only one reality, research is generalizable and the researcher themselves is an objective observer who has a limited effect on data and analysis (Higginbottom & Lauridsen, 2014). Instead Charmaz suggested a GT method that was underpinned by constructivism. Crotty (1998) describe constructivism as the assumption that meaning is generated by individuals as they interact with and interpret an object rather than a single meaning already existing within an object waiting to be uncovered. Therefore, constructivism conflicts with the ideology that there is a single objective reality that is amenable to being discovered through research. Higginbottom and Lauridsen (2014) describe Charmaz's arguments against positivism within GT using the constructivist stance. Firstly, the research and the researcher do not exist independently, meaning that the researcher's socially constructed reality informs their interpretations of the data. Therefore, the findings from research do not pertain to an objective reality but rather are interpretations of multiple of realities. Secondly, this is further compounded by the fact that participants also have their own socially constructed reality. The job of the researcher is to try and begin to discern under what conditions the participant's respective realities are constructed. Ultimately the findings from research are a tangle of the researcher and the participants socially constructed realities.

Based on these central tenets about the research process from the constructivist approach, Charmaz (2008) describes the four principles of CGT. Firstly, the research process should be treated as a social construction. This means that there are no definitive rules as to what a researcher should do and when they should do it. Instead Morse et al. (2016) argues that a flexible and dynamic approach to research should be taken. One that is inductive, comparative, emergent, and open ended by adopting the iterative logic of abductive reasoning. For example, when a researcher encounters a novel finding the study does not simply end there. The researcher should pursue it

further by (1) thinking of all possible theories that could explain the finding, (2) go back into the field to collect more data to test these theories, and (3) on this basis, determine the most likely theoretical explanation. The second principle described by Charmaz (2008) is critically examining research decisions and directions taken. Throughout the research process the researcher must ask themselves what they are doing and why they are doing it. According to Ramalho et al. (2015), using reflexive strategies the researcher ensures the 'groundedness' of theories by prioritising the data over their own personal preconceptions and/or prior knowledge of the subject area. Ramalho et al. (2015) describes the two primary reflexive strategies as the constant comparative method and memo writing. The constant comparative method is used to compare the existing literature with the data, codes, and categories to assess if the existing literature is relevant. Typically, the data should be prioritised over the existing literature. Memo writing encourages reflexive thought by providing an audit trail of thoughts, feeling and questions that emerge during analysis as well as the rationale for decisions made. Memos can also help the researcher with the delicate process of framing their research within existing literature without letting the existing literature dominate the findings. Charmaz (2008) refers to the third principle of CGT as the improvisation of methodological and analytical strategies during research process. Sengstok (2015) argues that CGT's focus on flexibility is an extension of Glaser and Strauss's original formulation of GT where the method was presented as a flexible and creative strategy that researchers could use as they saw fit. This flexibility allows the researcher to devise their own ways of *doing* GT. The researcher can choose from a variety of methodologies based on what bests suits the research area in question while still retaining the analytic skills from GT. Mills, Bonner and Francis (2006) describe how Glaser and Strauss aptly described GT as smorgasbord of techniques that researchers could choose from according to their tastes. In other words, a diligent CGT researcher will use a combination of techniques and methodologies that will transform their participants stories into theory. Charmaz's (2008) fourth principle of CGT is collecting quality data to enable the researcher to understand the world from the participants point of view. According to Lloyd (2007), this principle is the central tenet of CGT. Therefore, CGT is not preoccupied with the validation of theories or replicability that is associated with positivism. Instead CGT focuses on presenting findings that are congruous with the participants lived experience. To elucidate this point Lloyd (2007) refers to a comment made by one of their participants who says

'when will you tell us something we don't know'. This is exactly the aim of CGT, to present the findings in a such a way that the participants recognise their own experiences. Therefore, at the core of the CTG is the desire to give 'voice' to the participants. Morse et al. (2016) describes this process of entering the participant's world to try and get as close as possible to their reality of the situation. The aim therefore is to bring the participants perspectives to the forefront of the research process by gaining insight into their beliefs, purposes, actions, inactions, and motivations. CGT also attempts to discover the hidden implicit assumptions on which participants base their meanings and actions. In this way, the micro and macro levels of analysis are shown to be interdependent and the subjective and social can be interconnected.

1.7.5 SUITABILITY OF GROUNDED THEORY

GT was deemed to be the most appropriate method to use for both studies one *'When the lie is the truth: grounded theory analysis of an online support group for factitious disorder'* and two *'Claiming someone else's pain: a grounded theory analysis of online community user's experiences of munchausen by internet'* because they fit the criteria outlined by Gustavsson (2007). Furthermore, it was also decided to use CGT because it was methodologically congruent with both the characteristics of the research area and the philosophical stance of the researcher.

Gustavsson (2007) provides a list of four research situations for which Grounded Theory is ideally suited. It is argued that these research situations are consistent with the two studies of FD and MbI and are elucidated below:

(1) Phenomena lacking theoretical formulation – phenomena where there is a dearth of theories from an inductive and empirical standpoint. The existing theories surrounding FD and MbI both grapple with the absence of a solid evidence base. The incongruence between theory and reality is particularly apparent with FD, where clinicians struggle to diagnose and treat the disorder due to the ambiguous information available to them. Given the difficulty in accessing people with FD, the theories have been largely speculative and this may in part explain their lack of practical application value. Unique theories about MbI are non-existent due to limited research in this area.

Furthermore, what is known about MbI tends to be based on extrapolations from FD, which are speculative to begin with. The lack of substantiated theories around both FD and MbI make them ideal candidates for Grounded Theory research. Bhattacharya (2017) explains that this is because the goal of Grounded Theory research is to avoid a priori theories and begin analysis with an open mind. This means that any resulting theories should adequately capture the participants' experiences, actions, and interactions because they are firmly grounded in the data and not skewed by a priori assumptions.

(2) *Phenomena where nuances need to be emphasised* – capturing the complexity and variety in the interpretations of participants' experiences, actions, and interactions. Grounded Theory does not seek to develop a one size fits all theory but instead looks at the diversity in the data. This acknowledgement of the nuances in participants' experiences is particularly applicable to FD and MbI because first-hand accounts of the disorders are so scarce. As such it is not known whether there is a common experience of FD and MbI or whether the experience is more diverse and individual. First-hand accounts of other mental illnesses indicate that experiences are more likely to be diverse. While the basic nature of the underlying disorder remains relatively consistent the unique vantage point of the individual tends to colour their experience (LeCroy & Holschuh, 2012). Therefore, the ability to extract the nuanced experiences of both FD and MbI adds depth to the analysis and is an important advantage of using Grounded Theory.

(3) Where there is a need to make existing theories more relevant – the old theories no longer reflect reality and need to be updated to improve their relevancy. The current FD theories do not seem to offer any practical application value for clinicians. This indicates that there is a disconnect between theory of FD and the reality FD. Therefore, this thesis will use a new approach by using first-hand accounts rather than observational case studies to try and update the theories of FD, so that they reflect the lived experience of those with FD. Regarding MbI, existing theories are based on extrapolations from FD, which is in and of itself problematic. Furthermore, existing theories completely ignore the online dimension of the behaviour where MbI is played out. Therefore, first-hand accounts of MbI from those who encountered it will be used to develop updated theories that incorporate the experiences, actions, and interactions

of MbI within the online environment. The limited existing theories surrounding FD and MbI are outdated and require updating using first-hand accounts that acknowledge the lived experience.

(4) Phenomena that yield consequences – shed new light on a phenomenon that will have consequences for the actors and others involved in the area. The practical application value of the research conducted in this thesis is paramount. Particularly given that the current understanding of FD and MbI is of little benefit to the actors involved, including sufferers of the disorders, clinicians, and online community users. Using Glaser's (2009) approach to CGT, as these will be the first generation of studies using first-hand accounts, they will produce new hypotheses that will need further exploration and validation through additional qualitative and quantitative studies. These studies will not in and of themselves be verification of the theories but rather generate hypotheses that require further research. Congruent with Grounded Theory, while the goal is for these studies to have a practical application value, the studies mark the start rather than the end of the journey to achieving this goal. Ultimately any theories generated would require testing before being operationalised.

(5) As part of a larger study – using Grounded Theory can be effective as part of a larger study. The two qualitative studies in the thesis are essentially keeping in the tradition of the hypotheses-generating method of Grounded Theory with a twofold contribution. Firstly, the primary aim of the studies is to contribute to the larger study. That is in informing the development of a method to detect MbI and aiding in the selection and interpretation of features associated with MbI. Secondly, an important secondary benefit of collating information about FD and MbI is generating new theories based on first-hand accounts that can undergo further testing with the ultimate main of producing theories with a higher practical application value than the theories at present. Therefore, the two qualitative studies in this thesis are not only situated within the overall research aim of the thesis, but also within the larger domain of FD and MbI research. Situating studies within larger studies is congruent with the Grounded Theory approach.

Having decided that GT was suitable for the two studies, the next decision was choosing a school of GT. CGT was chosen based on several factors. Firstly, CGT

stresses the importance of giving 'voice' to the participants in such a way that they will recognise their experiences from the findings of the study. The primary aim of the two studies is to provide first-hand perspectives of FD and MbI rather than the previous studies which used observational accounts and had limited practical value. It is of paramount importance that the two studies adequately capture the lived experiences of those most affected by FD and MbI to ensure that the findings have value. Therefore, CGT is a good fit for both studies as its methodology seeks to uncover the lived experiences of the participants. Secondly, Charmaz (2006) emphasises the importance of collecting rich data, the most prized being written personal accounts as they offer the most novel findings. Both studies will be using personal accounts written in online forums from those who have FD or who have encountered MbI. These sources of data are consistent with CGT preference for thick descriptions from participants that will help reveal their feelings, views, intentions, and actions. Thirdly, Charmaz (2011) describes how CGT threads the fine line between application and innovation by providing a method for data analysis, that is absent in many other qualitative approaches, but not to the extent that the researcher does not have some creative flexibility to explore. This means that as analysis progresses the researchers can apply the method as they see fit to allow for novel findings to be pursued. Given that these studies will be the first to examine first-hand accounts of FD and MbI, it is important that there is some flexibility in the method to allow for previously unexplored ideas to percolate based on these new perspectives from participants. Fourthly, the ontological and epistemological underpinnings of CGT are consistent with the researchers own philosophical stance. That is a relativist ontology whereby there is no single truth but rather multiple human realities based on our own unique perspectives of real worlds that are multidimensional. A subjectivist epistemology whereby the construction of meaning is based in the interrelationship between the researcher and the participants. Therefore, the CGT methodology provides an ontological and epistemological fit with the researchers own position.

1.7.6 THREE CODING PHASES OF CONSTRUCTIVIST GROUNDED THEORY

According to Charmaz (2006) there are three phases of coding that bridge the gap between data collection and theory generation. These are (1) open coding, a code is

assigned to every line of text so an initial impression of the data can be formed; (2) focused coding, grouping the open codes according to contexts, consequences, causes and interactions to create categories; and (3) theoretical coding, hypotheses how the categories are connected to one another so they can be developed into a theory. A key tenet of the coding process is to stay open minded to all the possible theoretical directions of the data. To maintain this open mindedness Charmaz (2006) recommends constant critical examination and reflection on the data as well as identifying the meaning within the data. This is referred to as 'active coding' whereby the research continuously interacts with the data by asking numerous questions. Active coding allows for novel finding and helps to generate new research questions.

1.7.6.1 Open Coding – Identifying Codes and concepts

Strauss and Corbin (1998) define open coding as the analysis phase in which concepts, their properties and dimensions are uncovered. Open coding involves becoming immersed in the data by reading and re-reading the data several times over. At this stage labels can be attached to each line of data to represent what is occurring. When labelling, Charmaz (2006) suggests using gerund verbs which are the noun form of verbs. Using gerund verbs helps to emphasis actions and processes while maintaining the focus on the participants responses and contextual meanings. Charmaz's (2006) proposed guidelines for open coding including, remaining open to any theoretical possibilities within the data, staying close to the data to preserve the participants experiences, keeping the codes simple but analytical, using short codes that maintain the inherent action, constant comparison/revision of the codes throughout and moving through the data quickly to help facilitate spontaneity. The significant processes within the data were identified by being mindful of the following questions suggested by Charmaz (2006): What are the important processes? How can they be defined? How do the processes unfold? How do they participants act during the process? What do the participants claim to think and feel during the process? What might their behaviour indicate? When, why, and how does the process alter? and What are the outcomes of the process? The outcome of open coding is a variety of concepts that can be grouped together to create categories.

1.7.6.2 FOCUSED CODING – IDENTIFYING CATEGORIES

According to Charmaz (2006) open coding is complete when the analytical directions become apparent, the researcher can then move on to focused coding. Focused coding is the organisation of the codes according to frequency and significance. Decisions are made regarding which groupings of codes are analytically sensitive enough to be developed into categories. Through focused coding larger sections of the text can be explained. As Charmaz (2006) notes, focused coding means constantly going back over the data as issues that were previously implicit become explicit. Previous statements that were overlooked need to be re-examined as new avenues of analysis became apparent. This active involvement is the cornerstone of grounded theory, whereby data is acted upon rather than merely observed. Through action the researchers own preconceptions about the phenomenon under investigation can be challenged as new opportunities for analysis are revealed. Action involves comparing codes with codes to create focused codes and then comparing the data with the focused codes for clarification.

1.7.6.3 THEORETICAL CODING – IDENTIFICATION AND INTEGRATION OF CORE CATEGORIES

Charmaz (2014) describes theoretical coding as establishing the relationships between the categories identified during focused coding to develop higher level categories. Theoretical coding helps in reconstructing the data to create an analytical narrative that is comprehensive and comprehensible. Therefore, the aim of theoretical coding is not just describing the relationship between the categories but rather moving in a theoretical direction through the creation of an analytical narrative. To facilitate the development of an analytical narrative, Glaser (1978, 1998, 2005) has proposed a multitude of coding families that can be used to organise the categories. These include the 'Six C's' (causes, contexts, contingencies, consequences, covariances and conditions), process (e.g. transitions, passages, progressions), degree family (e.g. grades, continuum, levels), dimension family (e.g. sector, segment, part), type family (e.g. styles, classes, genre), identity-self family (e.g. self-image, self-concept, self-worth), cultural family (e.g. social norms, social values, social beliefs) and paired opposite family (overt-covert, in-out, informal-formal). However, Charmaz (2014) is sceptical about using these coding families for theoretical coding as they appear to be arbitrary and are far from comprehensive. Furthermore, there is a fine balance between

emergence and application whereby forcing data into these predefined families may be counterproductive for CGT where the emphasis is in letting the data guide the development of theory. Therefore, Charmaz (2014) suggests that researchers use these coding families as possible directions rather than being definitive coding families that data must be fitted into. Charmaz (2008) also warns against falling for the semblance of objectivity that accompanies theoretical codes. Theoretical codes are not objective evidence and are open to criticism. Therefore, while theoretical codes can be useful for honing analysis and critically evaluating if the codes adequately capture all the data, they should not be shoehorned into frameworks

1.7.6.4 QUALITY ASSURANCE

Charmaz (2006) argues that the quality of a grounded theory study should be judged based on four criteria, credibility, originality resonance and usefulness. Credibility refers to the extent to which the data supports the concepts, categories, arguments, and analysis. Originality is the reach of the study in challenging existing knowledge. Resonance is whether the participants can recognise their own experiences and can make sense of the findings. Usefulness is the practical contribution that the finding makes. The most weight is typically placed on the first criteria, credibility, as it provides the foundation for supporting the validity of the findings. One method of ensuring credibility is using inter-rater reliability or external auditor checks. However, a single coder is generally used to analyse the data. According to Levitt (2015) this is due to three inconsistencies between constructivist grounded theory and the assumptions underpinning the use of inter-rater reliability or external auditor checks. Firstly, the use of inter-rater reliability or external auditor checks are based on the realist premise that there is an objective reality that can be grasped through coding. However, for constructivist grounded theory reality is subjective and analysis will always be influenced by the researchers own understandings. Therefore, inter-rater reliability or external auditor checks only show that two researchers have learned to code in the same way rather than demonstrating that their coding is an accurate reflection of the phenomenon under investigation. Secondly, the use of inter-rater reliability and external auditor checks has the potential to weaken the analysis. Qualitative analysis involves the researcher becoming immersed in the study to the point that they gain an in depth understanding of the complexities and nuances involved. The authenticity and integrity of the study could be compromised if the

researcher is compelled to make changes to reach consensus with an investigator who does not share the same expertise from being immersed in the study. Therefore, according to the constructivist epistemology the researcher who is immersed in the data has the best insight because they have gone beyond a surface understanding of the phenomena under investigation and this helps to eradicate bias. Thirdly, the use of inter-rater reliability or external auditor checks is not compatible with the constructivist social justice ontology. The aim of qualitative analysis is to develop a theory that is trustworthy, compared to quantitative analysis where the aim is to produce a theory that can be replicated. According to Levitt (2015, p475) qualitative analysis seeks to shed light upon data that contain multiple meanings, contradictions, ambiguities, and subjective complexities and to create understandings that accurately represent these qualities'. Given the multiple perspectives of a single phenomenon, the same data set could produce different results that are equally valid depending on how the researcher situates themselves within the different perspectives of the participants. These findings can then be put to the test through further investigation in separate studies. The use of inter-rater reliability or external auditor checks is suited to a quantitative study where the aim is to develop a universal theory that requires consensus. However, for constructivist grounded theory these checks serve to limit the different understanding that are possible from the same data set based on multiple perspectives.

While the use of inter-rater reliability or external auditor checks are not generally suitable for constructivist grounded theory, other forms of quality assurance are possible. These include memo writing, inclusion of quotes and clear description of the analysis process (Birks & Mills, 2015; Sargeant, 2012; Corden & Sainsbury, 2006). Birks and Mills (2015) argue that memos are the cornerstone of quality for grounded theory by providing an audit trail of the decision processes that occurred throughout the study. They facilitate open communication by allowing others to view how the researcher engaged with the data and the series of thought processes that led them to their interpretation. The guidelines for writing memos provided by Charmaz (2008, p166) include (1) titling the memos; (2) writing memos throughout the study; (3) stating the properties of the code or category; (4) specifying the conditions under which the code or category emerges, is maintained, and changes; (5) comparing the code or category with other codes and categories; (6) including the data from which

the code or category has been obtained from; (7) describing the consequences of the code or category; (8) noting gaps in the data and opinions about it. While memos chronicle the decision processes that occurred throughout the study, quotes illustrate the categories and this lends credibility to the researcher's interpretations. A review of the quality assurance literature for qualitative studies by Corden and Sainsbury (2006) found that verbatim quotations were consistently cited as providing support for how evidence and conclusions were obtained. This is because verbatim quotes help to elucidate the connections between data, interpretation, and conclusion which helps provides validity, reliability, credibility, and auditability to the study. Another method of quality assurance described by Sargeant (2012) is a clear description of the analysis process to provide evidence of the trustworthiness of the analysis. Providing a description of how the different levels of coding are applied shows that the researcher engaged in systematic theoretical coding and the constant comparison necessary to identify the core concepts, their properties, and the relationships between them (Munhall, 2007). For example, Sbarani et al. (2011) provided a worked example of the coding conducted during their research study in table form as evidence of the rigorous application of the grounded theory method. Focusing on the rigorous application of the essential components of the grounded theory research method is argued to be the most practical measure of quality (Long & Johnson, 2000).

CHAPTER 2.0 WHEN THE LIE IS THE TRUTH: GROUNDED THEORY ANALYSIS OF AN ONLINE SUPPORT GROUP FOR FACTITIOUS DISORDER

Aideen Lawlor * & Dr Jurek Kirakowski

School of Applied Psychology,
University College Cork

Published: *Psychiatry, 218*(1-2), 209-218

Statement of contributions

Lawlor, A. (Candidate)
Writing and compilation of manuscript, established methodology, data collection, data analysis and data interpretation.

Kirakowski, J. (Supervisor)
Supervised development of work and editing of manuscript.

This chapter is a revised version of the journal paper referred to above

2.1 ABSTRACT

Factitious disorder (FD) is poorly understood because of the elusiveness of sufferers. What is known is based on speculation from observational case studies and this is evident by the manifold diagnostic and treatment issues associated with FD. This study sought to fill the gap in the literature and overcome the elusiveness of FD sufferers by analysing their text communications in two online communities. One hundred and twenty-four posts by 57 members amounting to approximately 38,000 words were analysed using grounded theory. The analysis showed that contrary to current theories of FD, motivation is conscious and not unconscious, members did experience symptoms associated with the disorder, and they were also upset by their behaviour and wanted to recover but were deterred by fear. Furthermore, a cursory comparison of the appetitive model by Orford (2001) with some of the characteristics of FD reported by a selection of members, tentatively indicates that there may be an association between FD and addiction. However, this is speculative and will require in depth research to establish if there is any weight to this superficial link between FD and addictive behaviour.

Keywords: Factitious disorder; Munchausen syndrome; Online support groups; Grounded theory; Addictive behaviour.

2.2 INTRODUCTION

The DSM-V divides Factitious Disorder into two types, FD imposed on self and FD imposed on another (previously FD by proxy). The diagnostic criteria for FD imposed on self are; (A) making up physical or psychological signs or symptoms or causing injury or disease with the deliberate intention to deceive; (B) pretending to be sick or injured or to be having problems functioning; (C) continuing with the deception, even without receiving any visible benefit or reward; and (D) behaviour is not better explained by another mental disorder, such as a delusional disorder or another psychotic disorder. The DSM-V states that the diagnosis primarily hinges on showing that the individual is covertly exaggerating, fabricating, simulating, or inducing symptoms of illness and injury. For example, manipulating a laboratory test by adding blood to a urine sample so an abnormality will be falsely indicated in the test. The reliance on establishing if symptoms are intentionally produced was also the primary method of diagnosis in the DSM-IV (Kanaan & Wessley, 2010). Clinicians collect circumstantial evidence to develop an index of suspicion and then confirm suspicion through irrefutable evidence of deception (Steel, 2009; McDermott et al., 2000; Meadows, 1982). Therefore, formal diagnosis only occurs when the patient either admits to feigning or is caught producing symptoms (Guzman & Correll, 2008).

2.2.1 REVISION OF DSM CRITERIA

The revision of FD within the DSM-V were in part brought about by the critique of the DSM-IV. The DSM-IV criteria for the diagnosis of FD were (A); the intentional production or feigning of physical or psychological signs and symptoms; (B) the motivation for the behaviour is to assume the sick role; and (C) the external incentives for the behaviour are absent. While some of the issues associated with the DSM-IV have been partially addressed by the new criteria used in the DSM-V there are still some outstanding. The debate around the diagnostic criteria for FD centres on overcoming the violation of nosologic principals including validity, reliability, usability, distinctive observable characteristics, and self-identification (Hamilton et al., 2009). Turner (2006) argued that criterion A of the DSM-V did not capture the essential features of the disorder and should be changed to 'lying or deliberate autobiographical falsification'. This allows for the inclusion of "pseudologia fantastica", voluntary false confessions, and impersonations and to distinguish

between self-harm and FD. While criterion A in the DSM-V still focuses on the production of symptoms, within the diagnostic features section of the manual there is recognition of the wide variety of ways that symptoms can be produced beyond simulation and induction by the inclusion of fabrication and exaggeration. The example given in the DSM-V is of a person pretending to be suicidal after the death of their spouse despite no death occurring or indeed the individual not having a spouse. The DSM-V acknowledges the subtler manifestations of FD highlighted by Turner (2006), but there is room for expansion to include more specific manifestations of FD including "pseudologia fantastica", voluntary false confessions and impersonations. However, Hamilton et al. (2009) argued that confining FD to the 'sick role' is justifiable because it is the aim of the behaviour regardless of means. While the DSM-V has rephrased 'sick role' to pretending to be sick, injured or having difficulty functioning, the underlying aim of the behaviour remains the same. Furthermore, Krahn et al. (2008) highlighted that the revisions suggested by Turner (2006) would not aid diagnosis as detecting and defining what is a lie or falsification is notoriously problematic. Turner (2006) also proposed that criterion B of the DSM-V should be descriptive and based on what is known 'the behaviour leads to, or is likely to lead to, self-harm' and criterion should be C removed. This revision would counteract the false premise of intentionality and incentives in the DSM-IV. As Turner (2006) argues that FD behaviour is not intentional because deceptively harming one's self does not follow common logic. Regarding incentive, the inference that motivation is based on external disincentives (self-harm) being perceived as internal incentives is unfounded. All that can be reliably observed is that sufferers self-harm despite the external disincentives. The DSM-V has not dealt with Turners (2006) issues in relation to criteria B and C. This may be because as Hamilton et al. (2009) and Krahn et al. (2008) argued, defining FD as self-harm excludes cases of FD where there is no self-harm or threat as well as those which occur outside of a medical context and it would also fail to distinguish Somatoform Disorder (SD) from FD. Although Turner's (2006) revisions are open to criticism, the crux of the argument that criteria should be based on what is known rather than inferred is an important one. Given the lack of direct research, it may be questioned whether the revisions in the DSM-V have been informed by evidence rather than inference.

2.2.2 DEPENDENCE ON CASE STUDIES

The relative rarity of direct research and consequent dependence on observational case studies has been widely lamented (Taylor & Hyler, 1993; Eastwood & Bisson, 2008). The dearth of such research is attributable to the elusiveness of FD sufferers who once confronted are known to strongly deny accusations of feigning even when presented with evidence of their deception (Pridmore, 2006). Of those who do admit to feigning very few will seek professional help making long term studies of FD difficult (Krahn et al., 2003). If they do enter treatment it is for a short period of time and they tend to be reluctant to open up about their experience (Eastwood & Bisson, 2008). The lack of first-hand accounts of the disorder means that basic information necessary for the formulation of diagnostic criteria including motivation and symptoms is scarce and as Turner (2006) highlighted is based on inference from observations.

2.2.3 MOTIVATION AND AETIOLOGY

The motivation to occupy the sick role is believed to be largely dependent on pre-disposing factors that result in the psychological deficits caused being balanced by enacting FD (Tasman & Mohr, 2011). These psychological deficits are rooted in early childhood trauma including physical and mental abuse (Sadock & Sadock, 2008). This early trauma is also associated with Personality Disorders (PD) and it has been suggested that there is strong co-morbidity between FD and PD, more specifically Borderline Personality Disorder (BPD) (O'Shea, 2003). Theories regarding the psychological deficits that motivate FD are abundant and largely unsubstantiated (Ford, 2010). They include, gratification of dependency needs, defence against psychosis, need for identity, need for mastery, internalised anger/masochism, learnt behaviour, problematic child/parent relationship, need for distraction and self-enhancement, displacement of rage, defence against loss, substitute for 'lost object', seeking sympathy/attention and using FD as a psychological coping mechanism (Ford, 1996; Maldonado, 2002; Dryer & Feldman, 2007). The problem with trying to determine the salient motivational factors underlying FD is that the research has been primarily observational. Even if it was possible to directly evaluate FD sufferers, the common view is that while FD sufferers intentionally produce symptoms their motivation to assume the sick role is unconscious (Feldman & Eisendrath, 1996). However, there is no evidence to support this distinction. In fact, Bass and Halligan

(2007) have argued that because making the distinction between medical (involuntary action, unconscious motivation) and non-medical deception (voluntary action, conscious motivation) are not attainable it should be excluded from the DSM. Instead the focus would be on the underlying psychiatric problems which precipitated the deceptive behaviour in the first place as opposed to trying to establish a diagnosis based on motivation which could be highly variable and impossible to establish.

2.2.4 SYMPTOMS

The internal subjective symptoms experienced by people with FD are largely unknown. There is a significant amount of information about the observable symptoms of the disorder, which are used to form an index of suspicion (Catalina et al., 2008). However, the question of internal symptoms tends to be ignored, because the motivation to assume the sick role is believed to be unconscious and therefore people with FD are perceived as being unable to self-identify. Hamilton et al. (2009), for example, claim that most people with FD do not express dissatisfaction with their deceptive behaviour, supporting the belief that people with FD are not perturbed by their behaviour. This lack of negative symptoms may explain the reluctance to seek help however it also may also be linked to the fear of losing trust by admitting to deception and the stigma associated with violating the social norms of the sick role (Pridmore, 2006; Hagglund, 2009). Of the few FD sufferers who enter treatment, there is no significant difference in outcome compared to those who receive no treatment (Eastwood & Bisson, 2008). This was attributed to sufferers not admitting to their behaviour or engaging in long-term treatment.

2.2.5 AIMS OF THE STUDY

There is a significant gap in FD literature regarding the lack of first-hand accounts of the disorder which has led to difficulties in reformulating the DSM criteria. Core questions such as whether the motivation to occupy the sick role is conscious or unconscious; whether symptoms are experienced; whether sufferers are disturbed by their behaviour, do they want to recover and if so why are they reluctant to engage in treatment have been left to speculation. The aim of the study is to help close this gap, by using a novel method to overcome the elusiveness of FD sufferers, through the analysis of their text communications in online communities for FD. The information

obtained from the analysis will be compared with previous FD research and will be used to develop new theories based on the accounts of FD from those directly suffering from the disorder as opposed to observational accounts.

2.3 METHOD

2.3.1 SAMPLE

Online support groups for FD were located using the Google search engine. The search terms used were 'factitious disorder', 'Munchausen syndrome', 'online support group' and 'discussion forum'. Two groups were identified, for Munchausen syndrome and factitious disorder. They were both established in 2002 and were located within a larger forum for various mental health problems. The groups cannot be regarded as active as there were only 30 new topics posted in 2011 in the Munchausen syndrome group and eight in the factitious disorder group. Although it is standard practice to use active online groups for analysis it was not possible in this study, as there were only two such groups. Both groups are completely accessible to the public. However, registration is required to participate in the group and view other member's profiles. Therefore, it was not possible to ascertain the participants' sociodemographic details. Both groups were moderated by the same person whose primary role was to provide emotional support.

2.3.2 LIMITATIONS OF SAMPLE

The current sample is not without its limitations. Members of the online support group are largely self-diagnosed as having FD as opposed to being formally diagnosed. Therefore, it could be questioned as to whether they genuinely did have FD. As formal diagnosis requires intention to assume the sick role to be established, the fact that members who were included in the study admitted to assuming the sick/victim role for emotional gains, means they most likely do have FD. However, the anonymity afforded by online support groups makes it impossible to assess the truthfulness of these posts. As all members are anonymous it is also impossible to assess if they were all unique individuals or whether there were 'sock puppets' among the sample pretending to be multiple members. This is further confounded by the association between FD and Munchausen by internet (Feldman, 2000). Whereby FD is enacted online, members with FD could have been creating a false version of their FD and

enacting it in the online support group in a double bluff. There is also the possibility that members used in this study are not representative of the general population as they may possess characteristics unique to those who use online support groups. Although the integrity of the sample and resulting posts can be criticised it is the largest collection of first-hand accounts of the disorder available because of the elusiveness of sufferers. On this basis alone it warrants analysis but with the implications of the sample limitations considered.

2.3.3 DATA COLLECTION

Between the two groups there were 170 topics. The study was only interested in the experiences of people with FD, therefore, topics and replies to topics were only included in the analysis if they related to the person's own experience of FD. For example, some topics and replies excluded were enquiries as to whether a relative has FD or offers of emotional support. In total, 124 posts by 57 members amounting to approximately 38,500 words were analysed. One member accounted for 15% of the data contributing 6000 words, nine members accounted for 44% contributing between 1000 and 4000 words and the remaining 41% was made up of 47 members who contributed 60–1000 words each.

2.3.4 ETHICAL CONSIDERATION

There are several special ethical considerations to be made when researching online communities and these have been outlined by Eysenbach (2001). Firstly, is the community private or public? As the communications within the group were viewable without registration, they were deemed to be public spaces. Secondly, is informed consent necessary? Informed consent was waived as the analysis was passive and the data was anonymised. The inclusion of exemplar quotes provided a challenge, as using verbatim quotes could result in the contributor's identity being exposed through a google search. Therefore, redacted quotes were used to provide a the 'thick description' necessary to exemplify the categories and for facilitating quality assurance (Jensen & Laurie, 2016). Despite the contributor's posts being available in the public domain and the use of direct quotes being ethically permissible, it is still important to protect the privacy of contributors. This is because at present there is a grey area as to what is best practice when it comes to online participant anonymity

(Saunders, Kitzinger & Kitzinger, 2015). Markham (2012) has argued that while people may be interacting in a public space, they may still hold an expectation of privacy. Therefore, a researcher must make their own ethical decision on a case-by-case basis. Given the repercussions for a person with FD were they to be identified, the decision was made to ensure their anonymity by redacting quotes.

2.3.5 DATA ANALYSIS

Constructivist grounded theory as described by Charmaz (2014) was used to analyse the data. This method of qualitative analysis lets the data guide the development of a theoretical framework rather than imposing one based on prior assumptions. This results in more novel findings. Given that what is currently known about FD is speculative and contested, using this blank slate approach was deemed appropriate. Furthermore, constructivist grounded theory focuses on the lived experience of participants, this is congruent with the current study as the aim is to capture the first-hand accounts of FD. Using the constructivist approach, the three stages of grounded theory analysis detailed by Charmaz (1990, 2006) were applied. These are open coding, focus coding and theoretical coding. A sample of the coding process for the core category 'Affection' within the theoretical category 'Motivation' is provided in Table 2.1 (See Appendix A for additional samples [1]).

During open coding, the posts from each member were grouped together and analysed independently. The flexible strategy for line-by-line coding as described by Charmaz (2006, p.50) was used to inform the labelling of each line of data. This included dividing the data into its constituent parts, defining the actions (using gerund verbs), uprooting implicit assumptions, deciphering inherent actions and meanings, clarifying the significance of the points, constant comparison of data with data and identifying any gaps. Charmaz (2006) stresses the importance of staying open minded during line-by-line coding so that new ideas can be discovered. Staying open minded proved to be particularly important for this study as during initial coding it became apparent that members experiences of FD contradicted the existing literature. Through constant comparison of data with data across all members the common processes and actions

[1] Verbatim text has been removed from Appendix A of the public version of the thesis to protect the privacy of the online community users.

were identified. It was often necessary to go over data repeatedly to investigate if these same processes and actions were also applicable to earlier respondent's experiences. For example, one member explicitly mentioned that FD was akin to addiction, this led to the revision of the preceding data where members had implied that FD shared characteristics with addiction but this had been missed on the first round of open coding.

Focus coding was begun once the comparison of data across members was not yielding any new processes and actions. The first round of focus coding was conducted on each of members' set of codes to identify the frequently occurring and significant codes that were theoretically similar within each of the member's groups of postings. This resulted in the development of very specific categories such as 'feeling ignored', 'feeling ashamed of wanting a genuine illness', 'FD eroding identity', 'feeling too embarrassed to seek help', 'wanting to recover' etc. The second round of focus coding limited the number of these unique categories so that the categories were either revised, reused, or reformulated. These new categories were then applied to the categories from the first round of focus coding to try and create more encompassing core categories. This process involved moving back and forth between the categories obtained from individual members so that they were constantly being compared and cross referenced. For example, the following sample of categories obtained from the first round of focus coding 'not being believed by psychologist', ''difficulty opening about FD', 'fearful seeking help is enacting FD', 'fearing repercussions if people find out', 'afraid genuine issue will not be believed', and 'difficulty in accessing information' were all included in the core category 'impediments to recovery' which was later renamed 'barriers to treatment'.

Once the core categories were refined, theoretical coding started. According to Charmaz (2006) the purpose of theoretical coding is to develop an analytical story by examining the relationships between the core categories. The core categories were cross referenced with one another to decipher common groupings that could provide an analytical narrative about an aspect of members experiences of FD. For example, the analytical story pertaining to treatment was brought together through the core categories 'barriers to treatment' and 'self-employed recovery strategies'. The hypothesis is that because FD sufferers experience internally and externally imposed

barriers to treatment they are forced to try and create their own recovery strategies to cope with FD. This iterative process of theoretical coding continued until relationships between all the core categories had been established to create theoretical categories and several hypotheses about FD based on members experiences. The discussion section provides an in-depth discussion of the theoretical categories, core categories, and the hypotheses generated from them, as well as their relationship to the existing literature.

Table 2.1 Sample of the coding process for the core category 'Affection' within the theoretical category 'Motivation'

Open Coding	Focused Coding Round One (Category)	Focused Coding Round Two (Core Category)
Behaviour motivated by unloved inner child Seeks missing elements from childhood from others now Believes behaviour is due to lack of love as a child Felt it was a given they would not be loved/wanted	Feeling that they did not receive love during their childhood and now falsely occupy the sick role as a way of feeling loved.	The speculated motivation of the FD sufferers need to falsely occupy the sick role was for affection. The form of affection received from occupying the sick role came in numerous guises, they all were ultimately motivated by a need for some form of affection
Questioning why they behave this way and similar others do not Wanting to know why they behave this way Not knowing why they crave attention/love Don't understand why they have overwhelming need for affection	Although recognizing that a need for affection drove them to falsely occupy the sick role, some had no idea where this 'need' came from and it appeared unconsciously motivated	
Emptiness due to isolation is filled by mental health professional Feeling heard and validated by psychologist Faked relapse if number of sessions with counsellor reduced Retaliating if counsellor suggested improvement Lying to maintain friendship in case they lose interest Lying to keep friends attention on them	Feeling socially isolated resulted in some falsely occupying the sick role because it was an aberrant means of maintaining/creating relationships with health professionals/friends that were fulfilling	

Crave the feeling of been cared for Never satisfied by nurturing given Wanting to be taken care of Lying to receive care from family Cutting was rooted in wanting to be cared for and comforted	Falsely occupying the sick role so they will be cared for
Feel loved/gratified via concern from loved ones after hospitalisation Enjoyed that college friends worried for them Lie to manipulate loved ones so they can feel special Exhilarating having people worry and stress about them	Enjoy the feeling of others being worried and concerned about them because it makes them feel special and valued
Pretended to have hand stuck in drain to be rescued Always fantasised about being rescued from rubble etc. Trying to resist their need to be rescued Pretending to be in constant crisis so they can be rescued	FD sufferers expressed their desire to be rescued from an emergency or some form of crisis

As part of the coding process Charmaz (2008) recommends the use of memos to provide a structure for exploring, investigating, and developing ideas about the data as well as interrogating these ideas. Memos also act as form of quality assurance to allow the bridge from data to theory generation. Therefore, throughout the study analytic memos were written to create an audit trail as well as providing an opportunity for reflexivity. A sample of memos is provided below in Table 2.2 and 2.3 (See Appendix C for additional samples). Memos written during line-by-line coding explored any potential analytical leads that were generated. The memos written during focused and theoretical coding were concerned with documenting the development of core categories and the shift towards theoretical categories and hypothesises generation.

Table 2.2 Sample of memos written during line-by-line coding

Sample memos
Childhood issues Across members there seems to be a number who describe having experienced childhood difficulties and many believe that the residual issues from their childhood were somehow related to their FD. The issues varied greatly and included, absent father, abusive mother, adoption, mental illness within the family and sexual/physical/mental abuse. Interestingly only one member describes being a victim of by proxy. There may also be a connection between early onset of FD and childhood difficulties that needs to be teased out more. The most common outcome of the childhood issues was feeling they had been emotionally abused and this resulted in an emotional 'void' most frequently associated with feeling unloved. It is speculated that the affection received from FD goes some way in filling this void. A few members describe the sensation having an 'unloved inner child'. *Tangle between genuine disorders/co-morbid/fake but now becoming real?* Members describe the various disorders they have experienced in addition to FD but caution is needed in separating out the different groupings within this. Some describe genuine disorders they have; however, these disorders need to be separated into being a cause of FD or a consequence of FD. Other members are more confused and state that what started off as a faked disorder has now become habitualised to the point that they feel it is genuine. This switch from something being feigned to becoming something real led to them being accused of 'crying wolf' by those close to them who were aware of their FD.

Table 2.3 Sample of memos written during memos written during focused and theoretical coding

Sample memos
Negative Impact of Behaviour on Others – doesn't capture the sentiment There is a large category 'Negative Impact of Behaviour on Others' that is too vague and needs revision. The category houses codes that refer to the detrimental effect feigning illness is having on those close to the sufferer. The person may not even be aware of FD, but they are still unwittingly providing care for someone who does not have the claimed disorder thereby putting them under unnecessary strain. Although this is at the crux of the category, it is the FD sufferers expressed guilt and shame around duping others into caring for them that provides the category with real depth and value. The shame and guilt experienced also contradicts previous theories about FD that the sufferers are simply oblivious. A category titled 'Negative Symptoms' might better capture the issues that members experience as a direct result of FD by including the categories isolation, poor quality of life, mental health issues, loss of real self, exclusion, addiction, low self-esteem within it. *Motivation – split/addiction* The motivation for enacting FD appears to be very varied and there are overlapping motivations. One distinction within the motivation category is motivations that are longstanding and those that simply act as fast catalyst and 'trigger' the need to immediately enact FD. There may be a relationship with addiction here, this has already been mentioned by several members, in that they felt a compulsion to enact FD even though they didn't want to, they felt compelled. In the same way triggers resulted in this urgent feeling to enact FD, even to the point that some were triggered by typically innocuous stimuli like medical soaps or seeing an ambulance.

2.4 RESULTS

There were four theoretical categories, which included the core categories, obtained from the grounded theory analysis. These were, characteristics of FD (early onset, feigned and genuine disorders, FD behaviour), motivation (long-term motivation, episodic triggers), symptoms and treatment (barriers to treatment, treatment strategies).

2.4.1 THE CHARACTERISTICS OF FD AMONG MEMBERS

The general characteristics of FD described by members included three core categories. Firstly, its early onset during childhood; secondly, the types of mental and physical disorders that were feigned as well as genuine co-morbid mental disorders; and thirdly, the behaviours associated with enacting FD.

2.4.1.1 EARLY ONSET

10% of members described their first episode of FD occurring between age 4 and early teens. They retrospectively rooted the origin of their FD behaviour to childhood events, for example, recounting jealously towards peers who were ill, wanting attention from a young age, making the connection between being sick/victim and receiving attention and a parental death triggering first episode of FD.

> *'I've been really thinking about why I do this. When I was little my sister had a seizure, I got jealous because I could see the attention she got from everyone, especially since she was the apple of my parents' eyes and I was ignored and emotionally abused. It also could be that when I was younger I used to fake being ill so I wouldn't have to go to school, but even then, I was so manipulative because I knew that I would have all my mum's attention'*

2.4.1.2 FEIGNED AND GENUINE DISORDERS

Table 2.4 displays the type of health issues mentioned by members that were feigned and the percentage out of all members who feigned them, as well as their genuine co-morbid mental health disorders.

Table 2.4 Feigned health issues and their genuine co-morbid mental health disorders as a % of all members.

Health issue	% of all members
Total mental (feigned)	23
Suicidal ideation	12
Schizophrenia	5
Panic attack	5
Eating disorder	4

Borderline personality disorder	4
Munchausen syndrome	4
Others [a]	11
Total physical (feigned)	26
Epilepsy/seizures	5
Diabetes	5
Physical injury	5
Concussion	4
Terminal illness	4
Others [b]	11
Total co-morbid (genuine)	19
Depression	11
Obsessive compulsive disorder	9
Borderline personality disorder	5
Eating disorder	5
Anxiety disorder	5
Self-harm	4
Bipolar disorder	4
Others [c]	7

a Amnesia, psychotic break, obsessive compulsive disorder, bipolar disorder, depression, dissociative disorder.
b Asthma, vomiting, appendicitis, fainting, muscle pain, cancer.
c Reactive attachment disorder, attention deficit hyperactive disorder, psychosis, Substance addiction.

2.4.1.3 FD BEHAVIOUR

A cohort of members described the conscious process of choosing an illness they wanted to feign and then researching it to make the portrayal as realistic as possible. In many instances, they reported fantasising about enacting FD and the affection they expected to receive.

'I wouldn't ever go as far as to faking cancer, at least I hope I wouldn't! But I do constantly fantasise about sickness for attention and about the deaths of loved ones so I can get attention and comfort from others. These "others" who shouldn't be

nearly as important to me as the family or friends I dream the deaths of'

A cohort of members claimed to employ a variety of methods to occupy the sick role. These can be divided into three categories, self-harm, exaggeration and lying. The primary goal of self-harm was hospitalisation where they would undergo minor medical procedures and in more extreme cases surgery. To warrant hospitalisation numerous forms of self-harm were used including cutting, aggravating a pre-existing illness, overdosing, ingesting poisonous substances, or otherwise purposely injuring themselves.

'I'm disgusted by what I'll do to get admitted into hospital, to look pale while I'm having a 'seizure' I put talcum powder on my face, concealer on my lips pale. I cut my tongue with a razor and then lie on the floor wetting myself while frothing blood from my mouth, it's disgusting. Last time when they were about to discharge me I stole a needle and injected air into my cannula so I would get sicker.'

Members also reported exaggerating pre-existing mental/physical health problems, for example, a headache was upgraded to a migraine and mild depression became suicidal ideation. Lying was motivationally distinct. It was not for hospitalisation but rather to maintain a relationship through pity by occupying the sick/victim role. The lies typically revolved around their life circumstances, such as: pretending there had been a death in the family, or they have been sexually abused. When the attention started to wane the lying escalated in severity, for example, pretending to have a terminal disease.

'I start a new friendship by being truthful, but then I get attached and want more. I start lying about myself and my life, like telling people I've been abused. When the attention wanes or the person gets sick of listening to the same stuff and I'm not moving on, the lies escalate. They lies come so easily and and still make sense and are realistic, like I've thought about them for weeks. It gets to the stage where it almost needs to be reported "I'm highly suicidal", so I continue the game until there is a danger I'll be found out, so I leave the person or make a recovery.'

2.4.2 MOTIVATION

Motivation was divided into two sub-categories long-term motivation and episodic triggers. The former refers to the constant underlying motivation behind FD and the latter were the immediate triggers which motivated episodes of FD. Table 2.5 displays the motivations for enacting FD reported as by members who discussed motivation.

Table 2.5 Motivations for enacting FD as a % of all members

Motivation	% of all members
Total long-term motivation	72
Affection	33
Difficult childhood	11
Coping mechanism	11
Enjoyment	9
Duping delight	9
Validation of genuine disorder	5
Identity issues	5
External gains	5
No reason for behaviour	2
Total episodic triggers	16
Emotional distress	5
Hospital stay for genuine illness	4
Life changing event	4
Emergency stimuli	4

2.4.2.1 LONG-TERM MOTIVATION

Approximately 72% of members directly and indirectly speculated on the possible long-term motivations underlying their need to occupy the sick role, these are discussed in detail below.

2.4.2.2 AFFECTION

Receiving affection was the most frequently cited motivation for occupying the sick role. Within this category 14% of all members reported feeling unloved, wanting to be cared for and cared about, wanting to feel special and rescued. For some, the social

isolation they experienced led them to occupy the sick role as a way of making contact with others, even if it was just a health care professional. Occupying the sick role was also used as a means of maintaining relationships with people they have become attached to.

> *'As soon as I'm alone I feel intense emptiness and loneliness and have the need to fill my life with somebody's love and attention, but of course as soon as I get a bit of attention I seek more and more. I've recently been alone and crave two people's attention particularly. I wish I could be part of their family and they would care for me all the time, so the best I can do is drag them in to lies which are completely unrealistic (getting bigger as the attention is thins) and I'm surprised they haven't realised I'm lying!'*

Members theorised that their pathological need for affection was rooted in a lack of love during childhood and they continue to try to fill this void by enacting FD. However, for the remaining 19% of all members, despite recognising that need for affection was at the root of their motivation, they were unable to give a reason for this need.

> *'What's wrong with me? Why can't I do the right thing? How did I get to this stage? I think if I understood this, if I could understand how and why it started, I will know how I can stop it. I find myself wanting to have something bad happen. I want to be raped, I want to be robbed, I want to be attacked and put in hospital, I want an abusive boyfriend... I don't understand. I want these things to happen so that the lies can be real.'*

2.4.2.3 DIFFICULT CHILDHOOD

Although three members did describe having a good childhood, the majority reported some form of abuse, predominantly mental abuse. This included lack of emotional support from parents, emotional abuse from mother, for example, put downs, unfavourable comparisons being made to absentee father or birth mother, hatred towards them and being a victim of by proxy. They were also either adopted or in foster care or their father was absented from the family. Regarding physical abuse, this included smacking, cutting, burning, and pushing. There were also instances of sexual

abuse reported. Other childhood factors mentioned included an unhappy and stressed mother, blended family from which they felt estranged, history of familial mental illness, irresponsible parenting, extra marital affairs, mother suffering from a chronic illness and childhood illness.

I believe my problems are due to emotional deprivation, its invisible and no one can see the hurt inside of me. I was adopted by parents with personality disorders. They adopted me privately as their application to adopt was denied. My adoptive mother revered my adopted brother and detested me. My adoptive brother became very sick and I saw how my mother doted over him, because I wasn't sick I believed I had no right to the love that I desperately wanted. My birth mother was an alcoholic and I was told I was exactly like her. My adoptive mother divided myself and my brother into him being good me being bad.

2.4.2.4 ENJOYMENT

On a more superficial level, members described the numerous ways in which they derived enjoyment from occupying the sick role. They enjoyed watching others concern for them and the attention that they received which they described as insatiable; the more attention they got the more they wanted it. The overall experience of being in a hospital setting was also enjoyable including the medical procedures. They also reported an adrenaline rush from enacting FD, from the drama surrounding the emergency of a life-threatening situation. Interestingly, for those whose primary motivation was enjoyment they did not want to be genuinely ill as they felt it would detract from their ability to enjoy the whole experience. In some instances, they described how they did not seek help for genuine illnesses, but only for the feigned ones.

'I think I do this because I like the "drama" of the emergency. The anaesthetists, nurse's doctors etc. I am scared and excited at once. My heart races and I'm breathless and dizzy. The nurses put pads on my chest for cardio version, the anaesthetists ask if I understand the risks, the doctors hand me consent forms to sign. It's all me, me, me to the point where I am almost hyperactive. I am looking forward to what is to come but still scared. Then the big moment is when the anaesthetist tells me they are putting me to sleep and that I will feel a little strange.

I have a routine, first I try to fight it, then I kiss a picture of my son, daughter and wife and keep a mental picture of them. If I die I want them to be the last thing I think of. Then I wake up and all the fuss is over and I start feeling depressed'

2.4.2.5 COPING MECHANISM

Whenever members experienced difficult life events or struggled with their emotions they described turning to FD as a coping mechanism. Although they recognised that it was an unhealthy coping mechanism they still relied on it and found it difficult to replace. In one instance, a user specifically referred to a means of coping with the lack of control they have over their own lives given by the sense of control that they feel from duping authority figures.

'It is funny controlling/manipulating doctors, a part of me wants to do that to "authority figures". Factitious Disorder is about control because if I want something from people I must control them because otherwise they would control me, unless it's an equal situation. This is why I struggle with therapists because they try to nurture instead of caring while remaining equal. I find ways to stop them being in control. ….every time this part of me gets near a therapist, I wind circles round them and usually succeed. Crazy I know, there is self-preservation, sadism and an inner child saying nobody loves me!'

2.4.2.6 SENSE OF ACHIEVEMENT (DUPING DELIGHT)

Some members reported that they felt a sense of satisfaction when duping doctors into believing the symptoms they are producing. For others, satisfaction was derived from setting out the goal of being formally diagnosed and achieving this diagnosis.

'When I am going to a doctor I exaggerate and I am vague when describing the illness in case they think I am not telling the truth. I get off on the diagnoses and attention. I often accuse the doctors of being stupid for not finding something very seriously wrong with me.'

'I tried faking schizophrenia, then I wanted borderline personality disorder. I have faked many types of mental illnesses but I never get a diagnosis of it and this makes me try even harder. Every time my medication is increased I feel like I have made

a HUGE achievement.'

2.4.2.7 VALIDATION OF GENUINE DISORDER

When it was felt that their genuine mental disorder was not being sufficiently acknowledged, members recounted how they either exaggerated accompanying physical ailments or produced symptoms, to receive the validation they felt was in proportion to their genuine symptoms.

'my main problem is Borderline Personality Disorder. I only fake physical symptoms when I am rejected or feel angry because I am not being seen or understood. I frequently feel that my real problems are not validated and that by acting out, it proves that it is real.'

2.4.2.8 IDENTITY ISSUES

Members reported feigning mental illness to provide an excuse for why they are the way they are, which they felt was as abnormal. For others, not occupying the sick role made them feel uncomfortable as being sick had become part of their identity. One member also described their desire to project a 'troubled' image of themselves and did so by feigning mental illness.

'I have faked panic attacks and said I'm highly suicidal. I am very ashamed but I have also said I have been abused. I have never been happy and have always lacked confidence, but I guess by saying I have experienced stuff that is so much worse than the truth, it gives me some sort of excuse for being the way I am.'

2.4.2.9 EXTERNAL GAINS

The most commonly cited secondary benefit of feigning illness was avoiding confrontations that were pending at work and receiving special dispensation.

'I continued to lie about small things, and even faked fainting and having a seizure to get out of work one day'

2.4.2.10 NO REASON FOR BEHAVIOUR

Of those who mentioned an underlying motivation, only one member could not offer

any explanation for their behaviour because from the outside they claim to appear fortunate and have few worries.

'I've done numerous things, the latest was pretending I had appendicitis, I got it removed and it was fantastic. I experienced a happy family life so I wonder why I cut myself and do things like this.'

2.4.3 EPISODIC TRIGGERS

Sixteen per cent of members described the circumstances and events which they believed tended to precipitate and hence trigger episodes of FD. These were either periods of emotional distress or environmental stimuli which they felt directly elicited the need to enact FD. Regarding periods of emotional distress, members described life changing events such as suicide, murder and divorce which preceded an episode of FD, feeling emotionally distressed when alone, escalated pre-existing fear of losing a loved one or being abandoned and not feeling in control.

'For me, I am scared of being abandoned so I frequently test people to check if they do care about me. I will push them to limits by giving them false reasons and eventually I do push them away and then reason with myself that I was right all along'

'When I'm not receiving attention from my friends I feel very lonely and start thinking about my pathetic life, I think this is my trigger for wanting attention because it helps fill my life and heart up. But I know this is so wrong.'

Members reported the emotional distress was further compounded when there was no one available to talk to and they reached out by enacting FD. Being in an environment where there were FD eliciting stimuli was also a trigger; this included stays in hospital for genuine health problems, watching medical related television programmes or seeing an ambulance or police officer on the street.

'if I hear an ambulance siren, see a police officer, or even watch ER or Law and Order SVU I start crave faking. Even if I see an emergency department, I will stop

and stare longing to be a patient. It's so difficult'

2.4.4 SYMPTOMS

Sixty-three per cent of members described the symptoms they were experiencing due to FD, particularly in relation to the negative side effects of having the disorder. Table 2.6 displays the percentage of all members who experienced the symptoms described in detail below.

Table 2.6 Symptoms experienced associated with FD as a % of all members

Symptom	% of all members
Total poor mental health	30
Suicidal Ideation	23
Depression	12
Suicide attempt	4
Total guilt and shame	28
Sense of guilt	23
Sense of shame	12
Total addiction	26
Compulsion	14
Need to enact FD overrides desire to stop	12
Uncontrollable	12
Addicted to enacting FD	12
Avoiding detection	9
Unconscious behaviour	5
Total dissociation from real self	26
Real/feigned symptoms merging	23
Erosion of identity	9
Total negative self-perception	25
Self-hatred	16
Dislikes behaviour	9
Believe they are innately bad	9
Believe they deserve punishment	5

Disgusted by thoughts/behaviour	5
Total isolation	19
Feeling alone in having FD	19
Total reduced quality of life	16
Hindering life goals	12
FD thoughts impacting quality of life	7
Incarcerated	4
Total ostracised	9
Not believed by therapist	5
Genuine illness not believed	5
Public stigma	4

2.4.4.1 POOR MENTAL HEALTH

Members expressed feeling depressed and hopeless due to FD and the resulting behaviour. In severe cases, they reported experiencing suicidal ideation and had attempted suicide as they felt they were in the desperate situation of simultaneously wanting to stop enacting FD but feeling that this is impossible. For those who did not attempt suicide the main deterrent was the effect it would have on their family.

> *'Rationally I know what I need to do but I somehow cannot stop myself. These powerful feelings are awful right now, I've genuinely thought about suicide, and I realise why some people drink and take drugs to the numb pain.'*

In one instance, a user experienced depression because they had disclosed their behaviour and had therefore relinquished the ability to occupy the sick/victim role, as opposed to the depression being a direct result of enacting FD.

> *'I am now on probation, if I fake again I will go back to prison for a long time. It's sort of a deterrent but god it's difficult. When you have been doing something all your life and then suddenly you have stop and have no coping mechanism in place of it, it gets very depressing.'*

2.4.4.2 DISSOCIATION FROM REAL SELF

Members discussed how their sense of whom they genuinely were had been lost, from hiding their FD from others and lying to occupy the sick/victim role. Some felt the ruse had become so ingrained that they were no longer able to distinguish between real symptoms and feigned symptoms. In extreme cases, they felt that the feigned symptoms were now genuine, for example, an individual feigning Obsessive Compulsive Disorder (OCD) had become habitualised into genuinely needing to carry out rituals which they had originally feigned.

> *'I researched mental illness and came up with depression, to make the depression worse I decided to cut myself as well. I cut into my leg and the next day showed the counsellor. Her concern was genuine and I can't explain how amazing it felt. I continued to research mental illness and made myself worse. Now the cutting was beginning to turn real. I was genuinely cutting myself to cope with depression and everything.'*

2.4.4.3 GUILT AND SHAME

Guilt and shame were the most frequently cited symptoms reported by members. As falsely occupying the sick role is considered taboo, members felt ashamed of their behaviour. Particularly in instances where they wanted to be physically or mentally ill and envied those who were genuinely ill. There was also a sense of guilt for lying to others, playing with their emotions and towards those who are genuinely ill. To cope with the guilt some strategies were often employed such as convincing themselves that the lie was true and by practicing dissociation when lying.

> *'I have to admit that the guilt and anger is the most difficult part. I don't think that will ever leave us because we are caring people. If we had no conscience there would be no guilt. A lot of people with this syndrome believe they are bad or evil, this just isn't the case, we were ill, that's it.'*

2.4.4.4 REDUCTION IN QUALITY OF LIFE

Members believed that FD consumed them at the expense of their quality of life. They became preoccupied with devising ways to enact FD or trying to resist the urge to enact it. FD hindered them from pursuing life goals either through careers not being

open to them, FD thoughts negatively impacting on their ability to function or being imprisoned for fraud due to their FD behaviour.

'The real problem is the thoughts because since I was in the hospital for a week I have felt the need to feed this stupid disorder that is part of my life… It's a daily struggle some days I just want to give up the only thing keeping me together is the work I do helping people... it's hard for me to accept there is no real help and that this is going be the rest of my life'

2.4.4.5 NEGATIVE SELF-PERCEPTION

There was a strong hatred towards themselves because of their FD behaviour, with numerous members feeling they were an innately bad person. They felt they should be punished for their FD behaviour: in one case, an individual was self-administering their own punishment through cutting. They were also disgusted at themselves at the lengths they will go to enact FD, for example cutting their mouth and wetting themselves to fake a seizure.

'At college, I became a known liar and while I didn't like it I felt I deserved it. The cutting was a way for me to express my anger towards myself, I was aware I had dug my own hole. I'm not suicidal, I hurt myself because it puts me in my place after my wrong doings.'

2.4.4.6 ADDICTION

The addiction like characteristics of FD were alluded to and directly referred to as an addiction by numerous members. The characteristics described included trying to avoid detection, describing behaviour as unconscious or uncontrollable, the need to enact FD being stronger than the desire to stop, having an overwhelming compulsion to enact FD, and the cyclical nature of their behaviour. That is, if FD is an addiction in a technical sense, members described it as such and perceived themselves to be addicted.

'My illness means I spend a lot of time in hospital and now I hate it. The thing is this is like an addiction. Sometimes you just need to score a hit. When a person stresses and worries over you, for some sick reason, it makes you feel fantastic'

2.4.4.7 ISOLATION

A cohort of members felt completely alone in suffering from FD before they came across the group and it was comforting for them to find similarities to others' experiences of FD. They reportedly did not realise they had a genuine mental health disorder but thought they were 'crazy'.

> *'I have read the posts multiple times because I can't believe the similarities. This is the first time in my life where I don't feel so alone. There are people out there like me who go to great lengths to feed this addiction. This post has changed everything for me, I want to get help and move forward'*

2.4.4.8 OSTRACISED

A cohort of members believed they were disliked by both the public and health professionals because they felt the disorder is not properly understood. In particularly they expressed frustration at what they perceived was the general misconception held by others that FD is a choice rather than a disorder. Some members also reported that once they had disclosed their disorder, health professionals dismissed them when they had a genuine illness.

> *'I got help for what I used to pretend, that I wanted to die and stuff like that but now it's real. I have no idea what to do, I hate myself for what I. I told my therapist that I really do want to kill myself now but she won't listen to me. My parents think it's not true and I don't know what to do. I really want to die but no one will listen to me.'*

2.4.5 TREATMENT

74% of members alluded to treatment under the two sub-categories, barriers to seeking support and recovery strategies. Barriers to seeking support refer to the 67% of members who described wanting to seek help, however they experienced obstacles that were both internally and externally imposed. The sub-category recovery strategies are the range of methods that members discussing recovery reportedly employed to try and recover from FD, in most cases members developed these strategies

themselves.

2.4.5.1 BARRIERS TO TREATMENT

19% of members described how they felt recovery from FD would require assistance. However, despite this only three members mentioned having sought professional help. They described their apprehensions about seeking professional support, with some even asking other members if it was possible to recover without it as their fear of seeking professional help was so great. Table 2.7 displays the number of members and the different barriers to seeking professional help they perceived.

Table 2.7 Perceived barriers to seeking help as a % of all members

Barriers to seeking support	% of all members
Fear of disclosing to therapist	16
Fear of losing family/friends	16
Fear of confronting issues	11
Logistical difficulties	11
Fear of recovery	5
Fear of perpetuating FD	4

2.4.5.1.1. FEAR OF DISCLOSING FD TO THERAPIST

Members felt they lacked the confidence and trust needed to tell a therapist about their behaviour and that once they had that the genuine symptom they experienced would not be believed due to a loss of trust. They also feared that they would be mocked by the therapist once they disclosed their problem. Three members reported that they had sought help for FD but were not believed by the therapist.

> *'A part of me would like to tell my therapist about this. However, I'm afraid that she won't ever trust me again, that she will question everything I tell her and that will make me more prone to acting out'*

2.4.5.1.2 FEAR OF CONFRONTING FD ISSUES

Members reported finding it difficult to admit their behaviour to others and themselves. By admitting it to others they felt they would no longer be able to enact

FD which had become a coping mechanism and furthermore they found it too emotionally distressing to talk about FD. Regarding admitting it to themselves, they described it as an overwhelming task to, and so despite being aware that their behaviour was problematic they preferred to consciously remain in denial.

'To recover, do I have to tell those I've lied to? I can't deal with the idea of admitting to others what I have done. I already don't like myself. I have caused others to worry, is telling them going to make it better? I can't talk to anyone about this, I really can't, not one to one. I want to ignore this, but I worry it won't just go way.'

2.4.5.1.3 Fear of losing family/friends

There was a perceived fear that if family and friends found out that they had been deceived and had been worrying unnecessarily, they would cease contact. This resulted in people with FD keeping their behaviour a secret and feeling they could not reach out for help.

'I'm very fearful of opening up to anybody about this, if my family and friends found out just imagine how hurt they'd feel that I have been lying to them'

2.4.5.1.4 Logistical difficulties

On a purely practical level, members reported experiencing problems in locating a professional who could treat FD. They described the long waits for appointments and found dealing with the health system was frustrating. They also found it difficult to find information about recovery options and some felt that this was because FD was poorly understood.

'when I was released from the psychiatric facility it took weeks to see anybody, this was an intake appointment for what program would suit me. Now I must wait another month for a second intake appointment. At that point, they may give me a counsellor to speak to. IM SO ANNOYED! I WANT HELP! It's not right, I am someone who really wants help to get better from this disorder and I can't get it!

2.4.5.1.5 FEAR OF RECOVERY

Members believed their behaviour was abnormal, however they were unsure if they genuinely wanted to recover due to the emotional needs the behaviour fulfilled. One user had an indifferent attitude towards recovery as they were content to stay in hospital rather than go through the recovery process from FD which they saw as being difficult for them.

> *'I've made the decision that I will go to a mental hospital and say I want to harm myself so they'll keep me in for a while. I don't know what I can do about this, I feel bad that some people are genuinely sick, I suppose therapy could help but I'd prefer to just live my life in a hospital.'*

2.4.5.1.6 FEAR OF PERPETUATING FD BY SEEKING HELP

Members reported the paradox as to whether seeking help for FD was a means of enacting FD. The fear that this was their primary motivation as opposed to wanting to genuinely recover made them feel apprehensive about seeking help, as it would only perpetuate their FD.

> *'I am thinking about getting a therapist, I'm starting to get that same sense of anxiety/excitement when I am think about creating a medical incident. I wonder if speaking to a therapist will just play into the issue. Doesn't that just feed it? Will speaking really help me? I have no idea what to do!'*

2.4.5.1.7 RECOVERY STRATEGIES

Members were aware their behaviour was abnormal and described wanting to stop enacting FD, primarily because they did not want to lie anymore and no longer wanted their lives to be dominated by the disorder. They also acknowledge how difficult recovery is and the poor prognosis of recovery; however, the majority were still hopeful that recovery was possible.

'Today I am starting to feel a bit different about myself, I think there is a chance of beating this. I know the prognosis for recovery is slight, but if you are truthful with yourself and want to change it must count for something'

To this end 21% of all members described the various strategies they had wittingly or unwittingly used which were successful. These are all displayed below in Table 2.8.

Table 2.8 Recovery strategies

Strategy	Description	No. of members
Therapy	Improved ability to function and lessened need for attention	3
Confronting precipitating issues	FD was a symptom of their underlying issues and was distracting them from dealing with these issues. Once they started to process the underlying issues the need to enact FD dissipated	3
Hero role	Helping others was found to negate the need to enact FD	3
Isolation	By isolating themselves from society they avoided triggers and did not have anyone to lie to	3
Changing environment	By immediately removing themselves from a situation where they have an urge to enact FD or taking respite from their normal environment where they are constantly battling FD urges	2
Fear of re-incarceration	Fear of returning to jail was their main deterrent (but they still struggled with their urge to enact FD daily)	2
Positive coping mechanism	Implementing positive	2

	coping mechanism to replace FD which was a negative coping mechanism	
Openness about FD	By disclosing their FD to others, they could no longer enact it	2
Self-insight	By recognising they had a genuine disorder and understanding the motivation behind it they could control their behaviour	2
Fear of being found out	By dwelling on the fear of being found out (but the concern that doctors had become suspicious was only a temporary deterrent)	2
New relationships	Making a concerted effort to form new relationships based on truth not lies	1
Communicating needs	Learning to communicate needs appropriately, instead of deceiving others to have needs met	1
Dependents	When a pet or person was dependent on them they logistically could not enact FD	1
Correcting lies immediately	When they lie they correct it straight away	1
Support system	Having the support of family and friends and someone they can contact	1

	24/7	
12 step programme	Applying 12 step programme to FD	1
Maintain boundaries	Endeavouring not to become overly attached to a person with whom they may enact FD	1
Need disappearing	No longer feel the need to enact FD, without having to take any conscious steps towards recovery	1
Positive changes	Trying to elevate mood by implementing positive changes to daily routine e.g. exercise	1
Logical thinking	When thinking logically can overcome the need to enact FD (but when emotionally vulnerable cannot maintain this logical state and enact FD)	1

2.5 DISCUSSION

The study aimed to fill the gap in the literature created by the lack of first-hand accounts of FD through the qualitative analysis of online support group member's interpretations of the experience of FD. The analysis led to 4 core categories and their respective subcategories being identified. These were 1. Characteristics of FD (early onset, feigned and genuine disorders, FD behaviour), 2. Motivation (long-term motivation, episodic triggers), 3. Symptoms and 4. Treatment (barriers to treatment, treatment strategies). The members capacity for self-reflection and spontaneous in-depth discussion of these topics indicated that motivation is not unconscious, while the motivations that they suggest are subjective interpretations, they do support previous theories and more importantly demonstrate the conscious nature of these

motivations. The analysis also challenged the perception that FD sufferers are not perturbed by their behaviour, do not experience internal symptoms and do not want to recover. Furthermore, the reasons for their reluctance to engage in treatment were obtained as well as the recovery strategies that they self-employ. A common thread running through all these categories and subcategories, which will be discussed in detail, is their overlap with addiction. On this basis, it is suggested that future research should examine whether FD is suitable for inclusion as a subcategory of 'Addictions and related disorders' as opposed to a standalone category in the DSM and whether this would aid clinicians in diagnosis and the treatment of FD.

2.5.1 MOTIVATION AND AETIOLOGY

2.5.1.1 SUBSTANTIATED MOTIVATIONAL THEORIES AND AETIOLOGY

There have been numerous unsubstantiated theories as to the motivation for assuming the sick role. The present study lends support to a number of these including, the need for affection, using FD as a coping mechanism, 'duping delight' (Ford, 2010), exerting control (Eisendrath, 1996), need for social interaction due to isolation (Savino & Fordtran, 2006), achievement of a sense of mastery (Feldman et al., 2001), poor sense of identity (Sadock & Sadock, 2008) and self-enhancement (Hamilton & Janata, 1997). An additional motivation theory was also obtained from the results, enjoyment. On a superficial level members reported the sense of enjoyment they derived from the hospital environment, undergoing medical procedures, the adrenaline rush of producing an emergency and the worry they caused, to the extent that some members did not want a genuine illness or would not seek help for a genuine illness because it would interfere with their ability to fully enjoy the experience.

The results also support the aetiology factors included in the DSM, which included family disruption (numerous members were either adopted or in foster care) and emotional and or/ physical abuse (the childhood abuse reported by members was predominantly emotional and inflicted by their mother) and a co-morbid PD, specifically Histrionic Personality Disorder (HPD) and Borderline Personality Disorder (BPD). Regarding the latter, only three members explicitly described themselves as having BPD however it was evident that this number is likely to be higher with the largest proportion of members reporting relational issues that

precipitated their need for affection being maladaptively met through the occupation of the sick role. For example, some members felt that the only way they could maintain relationships was through occupying the sick/victim role. Essentially, they created relationships that were based on pity, as they felt other-wise no one would be compelled to maintain a relationship with them. Other members described becoming overly attached to a specific individual and lying to them to receive increasing attention. In both cases there is a high possibility of co-morbid BPD. There was also evidence of co-morbid HPD, where those experiencing genuine mental health problems sometimes felt their suffering was not being given due recognition. Therefore, they would feign additional symptoms or exaggerate to elicit the attention/affection they felt was deserved for their original genuine mental health problems.

2.5.1.2 IS MOTIVATION CONSCIOUS OR UNCONSCIOUS?

The study found that 72% of members had spontaneously reflected upon their motivations for assuming the sick role. This is contrary to the common position held that while FD is the conscious production of symptoms it is unconsciously motivated. Reconceptualising FD as consciously motivated has significant implications for diagnosis. Firstly, motivation can be readily used to distinguish FD from both malingering and Somatoform Disorder (SD). The distinction between unconscious and conscious motivation is what traditionally separates FD from malingering (Eisendrath, 1996). However, it is still distinct from malingering, as it is for emotional as opposed to tangible gain and is beyond their volition. Regarding SD, the distinction has been difficult to make as both disorders are viewed as unconsciously motivated, by reconceptualising FD as being consciously motivated they can be distinguished from one another based on the criterion whether the motivation is unconscious or conscious. Furthermore, reconceptualising FD motivation as conscious as opposed to unconscious means that motivation can still be included in the DSM criteria, however this time with the inclusion of validated theories as to why patients assume the sick role as well as the internal symptoms experienced, instead of merely stating 'the motivation for the behaviour is to assume the sick role' which offers no guidance for diagnosis. This would contribute to making the FD criteria more nosologically sound.

2.5.1.3 CURSORY COMPARISON BETWEEN BEHAVIOURAL ADDICTION AND FD

A speculated link between the remaining theoretical categories is their relationship with the features of behavioural and substance addiction described by the excessive appetitive model by Orford (2001). It is cautiously suggested that FD may be an addiction to occupying the sick role. The excessive appetitive model seeks to explain the stages in which addiction is taken up, becomes established and is given up. According to Orford (2001) the primary processes involved in motivation for engaging in activities excessively are deterrence and restraint, and positive incentive learning mechanisms. Deterrence and restraint suggests that given the unfettered opportunity to indulge in pleasurable activities addiction is inevitable. Through positive incentive learning mechanisms, specifically positive reinforcement of the emotional rewards associated with the activity, the activity becomes addictive. Orford (2001) highlights that these emotional rewards are many and varied and will serve multiple emotional functions depending on the person's needs. This same motivational framework for developing addiction could be applied to occupying the sick role. In the case of FD the personal function of the behaviour was varied but in general it was aimed at resolving a perceived emotional deficit. To the extent that some members had become reliant on occupying the sick role as it was meeting their emotional needs whether it was a need for affection, as a coping mechanism, validating a genuine illness, resolving identity issues etc. The addictiveness of occupying the sick role could be further evidenced by the episodic triggers described by members including staying in hospital, medical programmes and seeing police and medical staff. This could be indicative of another primary process described by the excessive appetitive model, cognitive schemata where learning and memory are joined together. This is a cognitive bias in processing stimuli associated with previous occupations of the sick role to the extent that it triggers a desire to occupy the sick role again. In the following sections, the relationship between the characteristics of FD reported by members will be compared with those of addictive behaviour. However, it is important to note that while there may be some superficial links between FD and addictive behaviour, the weight of these links is at present tenuous. These links are speculative and require robust research to establish if they have any veracity.

2.5.2 SYMPTOMS

The internal symptoms experienced by FD sufferers were undocumented because of the perception that such symptoms simply did not exist as sufferers were oblivious to the disorder (Hamilton et al., 2009). The findings of this study contradict this assumption. Members did report experiencing symptoms associated with their awareness of the negative impact occupying the sick role was having on themselves and the people around them. These symptoms could be an example of the secondary amplifying process in the excessive appetitive model (Orford, 2001). Unlike the primary processes outlined, these are associated with negative as opposed to positive emotions. The secondary process creates cognitive dissonance whereby the negative emotions associated with the addictive activity can precipitate demoralisation through self-criticism, tension, confusion, depression, panic, self-attributions of bad character, guilt, shame and remorse, conflict with the emotional rewards produced through the primary processes.

A cohort of members also reported this secondary process whereby occupying the sick role produced negative emotions associated with demoralisation due to cognitive dissonance. The most frequently cited symptom was depression and in extreme cases suicidal ideation. This was primarily due to the conflict between wanting to stop occupying the sick role but not being able to. Depression and suicidal ideation are both well documented features of behavioural addictions including internet dependency, sex addiction and disordered gambling (National Research Council, 1999; Carnes, 2001; Kim et al., 2006). Members also described feeling shame and guilt due to their taboo behaviour, the hurt they caused others and their aberrant thoughts. These feelings of shame are also associated with taboo behavioural addictions, including sex and shopping addiction, whereby shame can perpetuate the addictive cycle (Adams & Robinson 2001; Yi & Baumgartner, 2011 Hall, 2012; Murali et al., 2012). Guilt is also a symptom of addiction, however contrary to shame it tends to be directed towards recovery, as a form of motivation (Dearing et al., 2005). Members reported feeling guilty about their thoughts and actions; for people suffering from sex addiction a study by Black et al. (1997) found that 61% felt guilty about their sexual thoughts, urges and behaviours and for gambling addiction, Chamberlain (2004) described that the guilt can be associated with risking their families' financial security through their actions.

While the specific factors underlying the shame and guilt differ according to the addictive behaviour involved, both the actions and thoughts associated with the addictive behaviour are universally a source of shame. In addition, some members explicitly stated the parallel between the symptoms they were experiencing and their perception of addiction i.e. trying to avoid detection, uncontrollable urge, urge to enact FD being stronger than the desire to stop, compulsion to enact FD and the cyclical nature of urges. These parallels are present in the four defining features of addiction as described in Ries et al. (2009), engaging in behaviour despite negative consequences, loss of self-control, behaviour is compulsive and experiencing urge/craving before engaging in behaviour. While the symptoms of FD appear to be outwardly like addictive behaviour, this is only on a superficial level and ultimately further research is required to assess if there is any substance to it.

2.5.3 BARRIERS TO SEEKING TREATMENT

FD sufferers are known to be reluctant to seek treatment however the reason for their elusiveness is unknown. It has been speculated that because motivation is unconscious and FD sufferers are unable to self-identify, they are simply not perturbed by their behaviour and do not seek treatment (Hamilton et al., 2009). However, members did describe wanting to seek treatment but were experiencing barriers that were largely self-imposed. Once again there are similarities to the barriers experienced by people with addictions who are reluctant to seek help. For members who were reluctant to seek help, it was driven by perceived fears, fear of disclosing to therapist, fear of loss of trust, fear of dealing with their underlying problems, fear of losing a coping mechanism, fear of losing family/friends and fear of recovery. These fears are also common among people with addictions, fear of disclosure (Weiss, 2004), fear of abandonment once the behavioural addiction is exposed Esterly and Neely (1997), fear of losing a coping mechanism (DiClemente, 2003), developing psychological defences to avoid confronting their fears around the issues which precipitated the addiction (Dayton, 2000) and the fear of loss of trust particularly in relation to addiction which has involved lying and secrecy e.g. sex addiction (Hall, 2012). While there are apparent similarities between the fears surrounding help seeking experienced by both FD sufferers and those with addictions, these may be specious and require further investigation to establish if the issues are grounded in a shared aetiology.

Additionally, there are two barriers to treatments reported by FD sufferers which may be unique. The first unique barrier is logistical difficulties, finding information and professionals who can assist them in recovery. The logistical difficulties experienced could be related to the fact that FD is poorly understood due to the reluctance of FD sufferers to participate in exploratory analysis (Sadock & Sadock, 2008). Therefore, very few professionals specialise in FD because of the lack of information available and because the demand for treatment is low. The second unique barrier is the fear that seeking treatment for FD may be another means of enacting FD, that they would use seeking treatment for their legitimate disorder as a way of also enacting the disorder.

2.5.4 RECOVERY STRATEGIES

Only four members reported using structured therapy as a recovery strategy the remainder were using a variety of methods to self-manage. This equates to what is referred to as untreated recovery in addiction research in which there tends to be low rates of treatment seeking and high rates of self-management (Slutske, 2006). According to Orford's (2001) model of excessive appetite self-management strategies are arguably more successful because they are precipitated by the person coming around to the idea that their behaviour is morally aberrant either socially or personally and enacting change for themselves.

As with barriers to seeking treatment, there was an overlap with the common self-management strategies employed in addiction and those employed by FD sufferers including natural remission (needs disappearing), avoiding triggers (isolation/changing environment/maintain boundaries), social and family support (support system, openness about FD, communicating needs), cognitive appraisal (logical thinking, confronting precipitating issues, self-insight), creating a positive milieu (positive changes, new relationships, correcting lies), avoiding jail (fear of re-incarceration, fear of being found out), adopting new coping skills (hero role, positive coping mechanisms) and responsibilities (dependents) (McIntosh & McKeganey, 2002; Sobel, 2007 Van Wormer & Thyer, 2009; Toneatto, 2013). However, the actual success of these of these strategies is unknown. While the coping strategies associated with FD and addiction may appear alike, it is at a superficial level and is not evidence

of a shared aetiology.

2.6 LIMITATIONS

The current study is not without its limitations. The similarities between the features of addiction and FD are at this point purely speculative. This is because the similarities have only be noted at a superficial level of analysis making it difficult to discern if they are truly indicative of a genuine shared aetiology between FD and addiction. Ultimately further research is needed to explore if the relationship between the shared features of FD and addiction is based on a common aetiology. Members of the online support group were also largely self-diagnosed as having FD as opposed to being formally diagnosed. Therefore, it could be questioned as to whether they genuinely did have FD. As formal diagnosis requires intention to assume the sick role to be established, the fact that members who were included in the study admitted to assuming the sick/victim role for emotional gains, means they most likely do have FD. There is also the possibility that members used in this study are the exception as opposed to the rule. Because they were posting on an online community they may be more aware of their behaviour and motivations for it and more inclined to want to recover. Therefore, members of the online community may not be representative of most people with FD because of their increased level of awareness. This could account for the discrepancy between the traditional view of FD and the one presented in this study. However, the traditional conception was based on speculation and individual case studies. The current study offers an improvement by using first-hand accounts and is a relatively large-scale study in this area. Despite the sample appearing biased it does reflect the undisputed characteristic associated with FD elusiveness, only three members sought professional help. But rather than interpreting this elusiveness as apathy or ignorance of FD, by having access to first-hand accounts the current study reinterprets this as fear of disclosure.

2.7 CONCLUSION

Theories regarding the aetiology of FD are speculative, based on observational case studies. This is reflected in the practical diagnostic and treatment problems which surround the disorder. The debate has become stagnated because of the limited information available on what it is like to have the disorder from the sufferer's

perspective as FD sufferers are highly elusive. The current study used a novel method of accessing first-hand experiences, analysis of text communication in two online communities for FD. The study set out to examine several core questions about the nature of the disorder: is the motivation to occupy the sick role conscious or unconscious, are symptoms experienced, if so what are they, are sufferers disturbed by their behaviour, if so do they want to recover and why are they reluctant to engage in treatment? The analysis found members were aware of their motivations, challenging the common perception that motivation is unconscious. Furthermore, members provided numerous theories as to why they felt the need to occupy the sick role, while these are admittedly interpretive, the majority concur with previous theories. Contrary to the conventional view of FD, members also described experiencing internal symptoms associated with the disorder, reported being upset by their behaviour, and claimed to want to recover. However, they were deterred from seeking formal help by anticipated fears and instead opted for self-management strategies. It is speculated that the experiences of FD reported by some members may share similarities with addictive behaviour based on the model of excessive appetite described by Orford (2001) and direct addiction research. They superficially appear to share the same primary and secondary processes and both groups are reluctant to seek formal treatment instead preferring self-management strategies. However, this tentative link between FD and addiction comes with a caveat as the comparison is cursory and superficial. Nonetheless, given the potential benefits of conceptualising FD as an addictive behaviour, for both the diagnosis and management of the disorder, it is important these superficial similarities are highlighted to stimulate future research.

CHAPTER 3.0 CLAIMING SOMEONE ELSE'S PAIN: A GROUNDED THEORY ANALYSIS OF ONLINE COMMUNITY USER'S EXPERIENCES OF MUNCHAUSEN BY INTERNET.

Aideen Lawlor * & Dr Jurek Kirakowski

School of Applied Psychology,
University College Cork

Submitted for peer review in *Computers in Human Behaviour*

Statement of contributions

Lawlor, A. (Candidate)
Writing and compilation of manuscript, established methodology, data collection, data analysis and data interpretation.

Kirakowski, J. (Supervisor)
Supervised development of work and editing of manuscript.

This chapter is a revised version of the journal paper referred to above

3.1 ABSTRACT

Munchausen by Internet (MbI) became increasingly visible in 2016 after several high-profile cases were reported in the media. It has not however received significant research interest. What is known about MbI is dependent on a series of case studies by Feldman (2000). The present study seeks to add a new perspective on MbI by analysing the experiences of respondents with an intimate knowledge of the condition. Posts from 12 online communities by 556 respondents discussing MbI were collected, amounting to approximately 91,300 words. These were analysed using grounded theory. The findings show that MbI perpetrators appear to advantage themselves by occupying 'ideal victim' personae. Despite this supposed sinister premeditation and the resultant emotional distress caused, an unexpected proportion of respondents were sympathetic towards the MbI perpetrators. However, the presence or possible presence of MbI eroded trust amongst participants and led to more reserved participation to prevent being duped or labelled as MbI perpetrators themselves. Both these strategies undermine the therapeutic benefits of online communities by restricting opportunities to confer normality and cultivate interpersonal support. There are no definitive clues available to identify the behaviour and no deterrents to discourage it. New innovative methods are needed to tackle MbI and several tentative suggestions are made including a text classifier, reverse photo search filter and a more onerous membership screening processes.

Key words: Munchausen by Internet; Online communities; Factitious Disorder; Online support groups; Efficacy of online support group.

3.2 INTRODUCTION

Successful online communities are built on interpersonal trust which ensures that participants feel invested in a community to the extent that they actively participate (Lambropoulos, 2013). One threat to interpersonal trust is identity deception which undermines the integrity of an online community (Dahlberg, 2001). Donath (1998) argued that identity deception, or even the possibility of identity deception, results in participants being less trusting of new members and less tolerant of those who do not adhere to the community's norms. Munchausen by Internet (MbI) is one such form of identity deception where a participant creates a persona eliciting sympathy. Despite the potential of MbI to erode interpersonal trust in online communities, it has received little research attention compared to other forms of identity deception. This study was undertaken to gain insight into MbI from the perspective of participants in online communities who have encountered MbI, and to identify new avenues for future research.

3.2.1 ONLINE COMMUNITIES AND TRUST

Belonging to an online community has been shown to offer a myriad of benefits which may explain the growing popularity of such communities. These include offering various forms of support including informational, emotional, esteem, network and tangible assistance (Coulson, 2005; Eastin & LaRose, 2005; Hwang et al., 2010; Idriss, Kvedar & Watson, 2009; Mo & Coulson, 2008; Obst & Stafurik, 2010; Setoyama, Yamazaki & Namayama, 2011). Research tentatively suggest that the availability of these supports in online communities helps to improve the general wellbeing of participants (Kummervold et al., 2002; Miyata, 2008; Nimrod, 2012, 2013; Rains & Young, 2009; Rodgers & Chen, 2006). However, the potential benefit of participation in online communities is dependent upon cultivating interpersonal trust, as it a prerequisite for initial and continuous participation, particularly where personal issues are shared (Benlian & Hess, 2014; Preece & Maloney-Krichmar, 2003; Ridings, Gefen & Arinze, 2002). Ridings et al. (2002) found that interpersonal trust in online communities is made up of two dimensions, ability (accuracy, soundness and reliability of the information shared) and integrity/benevolence (reciprocating and adhering to the norms of the community). However, as Donath (1999) noted, given the disembodied world of online communities, these forms of interpersonal trust can

be easily violated by identity deception. The presence of MbI throws the ability and integrity/benevolence of all participants into question and threatens the potential benefits of participation in an online community.

3.2.2 Online Communities and Identity Deception

Jeter (2015) describes three types of motivation for engaging in online identity deception including privacy concerns, expression of the true self/ or exploring identity, and mental illness or mood disorders. Identity deception for reasons of privacy and expression/exploration is often viewed as legitimate and benign (Bowker & Tuffin, 2006). In contrast, identity deception motivated by emotional and mood disorders is viewed more negatively as it can have malicious outcomes, intended or unintended. One type of identity deception that falls under this category is identity replacement. Segovia and Bailenson (2013) define identity replacement as the act of replacing one's own identity with another, either that of a real person or a fictitious identity. It is highly complex as participants must hide their real identity while simultaneously convincing others of their fake identity by adopting a whole new persona. The motivation for engaging in identity replacement is difficult to ascertain as there have been few studies that directly assess the perpetrators and victims. There are likely to be categories of identity replacement that have their own unique motivations and underlying psychopathologies. One such category that has been discerned from research by Feldman (2000) is MbI which involves falsely occupying the sick/victim or hero role.

3.2.3 Munchausen by Internet

MbI is the online equivalent of Factitious Disorder (FD). FD is a mental disorder where a person purposely feigns an illness and/or induces the symptoms of an illness to occupy the sick role (American Psychiatric Association, 2013). Similarly, MbI typically involves fabricating a fictitious identity online which revolves around personal illness and crisis (Feldman, 2000). It is not known if MbI is a standalone disorder independent of FD or whether those with FD are simply expressing their disorder online in the form of MbI. Given that MbI perpetrators have yet to be directly studied the only option so far has been to base the aetiology of MbI on extrapolations from FD based on their apparent similarities. However, it is important to note that applying such extrapolations is problematic because much of what is known about FD

is speculative and unsubstantiated given the elusiveness of sufferers (Ford, 2010; Lawlor & Kirakowski, 2014).

The primary purpose of FD is believed to be the fulfilment of psychological deficits due to pre-disposing factors including a poor sense of self, unmet dependency needs and childhood trauma (Fleming & Eisendrath, 2011; Tasman & Mohr, 2011). The corresponding psychological benefits of occupying the sick/victim/hero role which fulfil these deficits are attention, caring, nurturing, love, control, sense of importance and developing a sense of identity (Ayyer & DeSousa, 2014; Fleming & Eisendrath, 2011; Lasher & Sheridan, 2004). Based on observational case studies, traditional theories state that those with FD are oblivious to their aberrant behaviour, motivation is unconscious, no adverse symptoms are experienced, perpetrators are unperturbed by their deceptive behavior and therefore have no reason to seek help for recovery (Feldman & Eisendrath, 1996; Hamilton, Feldman & Janata, 2009; Krahn, Li & O'Connor, 2003). However, research by Lawlor and Kirakowski (2014) has challenged these traditional theories by studying first-hand accounts of FD, where participants were found to be aware of possible motivations for their behaviour, experienced a range of negative symptoms, were upset by their behaviour, and wanted to recover but were deterred by fear, resulting in a reliance on self-recovery strategies. Those with MbI may share the same characteristics as FD, however until such time that direct research is conducted extrapolations made from FD are speculative.

Despite the potential for MbI to erode interpersonal trust, thus compromising the benefits of online community participation, MbI has also received little research attention with regards to its impact on the functioning and dynamics of online communities. There has only been one study to date that directly investigated MbI. Feldman (2000) reported on the fallout in online communities when a case of MbI is exposed and outlined the cues associated with the behaviour, based on four case studies. Feldman (2000) found that the reaction in the online community when MbI is exposed was mixed. Some participants were perturbed by the hoax and tried to contact the MbI perpetrator to express their sadness and anger. In some instances, participants wanted to sabotage the MbI perpetrators' real-world identities by making their offline acquaintances aware of their deceptive behaviour. Others were unperturbed and found the whole ruse amusing, particularly for the audacity involved. Exposure of hoaxes

also led to divisions within the online community, those who believed it was a fake and those who believed it wasn't. Those participants who were disgusted by the whole experience tended to leave the online community, leaving those who stayed to process their emotions. Although not addressed in Feldman's (2000) study, undetected cases of MbI may also have the potential to cause disruption, as MbI perpetrators have no direct experience of the roles being falsely occupied. MbI perpetrators are also likely to spread misinformation by discussing their fictitious personal histories, medical advice received, medical tests they have undergone, disease progression and treatment options/outcomes. Given that the information shared in online communities is used when making health care decisions, following the misinformation spread by a MbI perpetrator could have significant negative consequences (Mo & Coulson, 2014). In fact, Witney, Hendricks and Cope (2015) have suggested that online communities should employ health professionals who can correct any misinformation provided by participants. Undetected MbI also has the potential to impede the recovery of genuine participants. Sharing positive stories of recovery is central to the success of support groups as it gives other participants a sense of hope which helps promote recovery (Kurtz, 2015). In an online support group where MbI is widespread or an MbI perpetrator dominates the community, positive stories of recovery will not be forthcoming as the goal is to receive attention which necessitates the use of dramatic/ a-typical stories which never progress to recovery. Based on observations Feldman (2000) also listed the various cues he believed were indicative of MbI, these include:

- Copies from others, text books and health related websites.
- Posting behaviour is not reflective of the severity of illness being portrayed.
- Characteristics of the health issues and treatments are caricature like due to misconceptions.
- Oscillating between being close to death and miraculous recoveries.
- Making fantastic personal claims which are contradicted or disproved.
- Constant flow of dramatic events which intensify as attention wanes.
- Claims other participants' inattentiveness is compromising their health.
- Avoiding telephone contact.
- Feigned blitheness about a crisis.
- Using 'sock puppets' with a similar writing style.

However, Feldman (2000) warns that MbI is particularly difficult to identify based on observational cues as there is often a mix of facts and lies. If it is successfully identified, punitive action beyond banning is rarely taken. Two options which have been suggested are: publicly exposing the identities of perpetrators; and legally pursuing perpetrators via their IP addresses, by treating the behaviour as a cybercrime (Feldman & Peychers, 2007; Pulman & Taylor, 2012).

What is known about MbI is based primarily on Feldman's study (Feldman, 2000) conducted over 15 years ago. Since then it has received little research attention despite the deleterious impact that identity deception has on interpersonal trust and consequently the therapeutic benefits of online community participation. Anecdotally it does appear to be a significant issue with high profile cases of MbI being reported in the media frequently including a teenager posing as the mother of a terminally ill child, a person pretending to have leukaemia and falling into a coma, the blog of an immigrant in London with terminal tuberculosis which turned out to be a fake and a wellness blogger who falsely claimed to have cancer (Jenkins, 2012; Kleeman, 2011; Montague, 2015). However, one of the most elaborate and infamous cases of MbI was the 'Warrior Eli blog' that ran for a decade. The fake blog chronicled John and Emily Dirr's son's experience with cancer. 71 fake online profiles were set up purporting to be friends and relatives of the Dirr's who communicated with one another to try and give the story credibility. Participants became suspicious because of an implausible stream of tragedies that seemed to have occurred to the family, including Emily Dirr's death in a car crash on Mother's Day, where she miraculously gave birth to the family's eleventh child in the moments before her death. Wright (2013) exposed the ruse and the pictures of the Dirrs and their 71 friends and relatives were found to be stolen from social media accounts. The perpetrator turned out to be a 22-year-old female training to be a doctor.

While these cases provide some glimpse into MbI the aim of the present study is to go beyond observations and analyse the first-hand experiences of a large sample of online community participants who have directly confronted MbI. This will help in gaining a broader perspective of MbI from the viewpoint of those in contact with MbI perpetrators, in better understanding the issues surrounding the behaviour and its

impact on participants and their communities and in stimulating further research into MbI.

3.3 METHOD

3.3.1 DATA COLLECTION

Threads within online communities pertaining to MbI were identified using the Google search engine. A variety of broad search terms were used followed by more specific search terms. The broad search terms included 'online community', ‘discussion forum', 'message board' and 'online support group'. As the term MbI is not widely known, characteristics which are associated with the behavior were used to create specific search terms to identify relevant threads within online communities. The specific search terms included, 'troll', 'fake', 'catfish', 'imposter', 'Munchausen', 'hoax', 'liar', 'attention seeker', 'alert', 'banned', 'false' and 'pretend'. The relevancy of the threads identified were assessed on the basis that a person(s) who was suspected of or confirmed to be falsely occupying the sick/victim and/or hero role was being discussed. Several criteria were used for identifying such threads where a genuine case of MbI was being discussed rather than a case of trolling. These were where the member under discussion had been explicitly banned by a moderator for MbI behavior, the member admitted to being an MbI perpetrator and where members had conducted their own investigations and collected direct evidence to prove the member was an MbI perpetrator. Direct evidence included stolen pictures that the MbI perpetrator was passing off as their own, no obituary notices appearing when an MbI perpetrator claimed a loved one had died, multiple accounts sharing the same IP address that was traced to the MbI perpetrator, no public record of the dramatic life events the MbI perpetrator was claiming to experience e.g. a car crash that does not appear in the news, and the person being portrayed online by the MbI perpetrator not existing offline when other members try to make contact e.g. no record of the person being portrayed online at the workplace they claimed to be employed at or no record of them at the hospital they claim to be in.

Using this method, 12 online communities containing threads where members discussed a case of MbI were identified. Six of the online communities were centered on parenting, the remainder centered on finance, incarceration, nursing and health. The

threads from the 12 online communities were collected, containing the posts of 556 respondents amounting to approximately 91,300 words. One respondent who was subsequently exposed as an MbI was deleted from the data set. To facilitate data analysis, each respondent was allocated an alphabetical code which was color coded according to the online community the respondent belonged to.

3.3.2 ETHICAL CONSIDERATION

The ethical guidelines for internet-mediated research as outlined by the British Psychological Society (BPS) (2013) were followed in conducting this research. The main issues pertained to privacy and consent. With regards to privacy, only data from online communities in the public domain was collected. However, this does not automatically imbue consent. Consent from the respondents in this study would be difficult to obtain given the large number of cases and logistical issues. For example, some of the respondents were no longer active, were anonymous 'guests' or did not have contact details. Given the contribution this study will make in understanding MbI, safeguards that were put in place to protect the anonymity of the respondents and given the passive nature of the analysis, consent was waived.

The safeguards put in place to protect the anonymity of the respondents included redacting the respondent's username, online community and verbatim quotes that could be used to identify the respondent. The redacted quotes included in the results section provide a 'thick description' that serves to both illustrate the categories and provide quality assurance (Jensen & Laurie, 2016). It is important however that verbatim quotes are not included as they could be used in a search engine to identify the respondent. The effort made in ensuring the privacy of the participants is important given the ongoing debate around online user's concerns and expectations of privacy and the ease by which a person's real-world identity can be exposed through jigsaw identification (Saunders, Kitzinger & Kitzinger, 2015). Furthermore, given the personal nature of the information being collected it would be ethically irresponsible to compromise the respondent's anonymity by using direct quotes. While according to the formal ethical guidelines direct quotes could be used, Markham (2012, p.4) succinctly sums up the ethical dilemma for the online researcher *'what may seem ephemeral or innocuous at one point in time might shift rapidly into something that*

causes real or perceived harm'. Therefore, it is better to err on the side of caution and take measures to protect the anonymity of participants.

3.3.3 DATA ANALYSIS

Constructivist grounded theory as described by Charmaz (2014) was used to analyse the data. MbI lacks substantiated pre-existing theories meaning it is well suited to the CGT method. This is because CGT discourages the reliance on pre-existing theories and instead advocates an open mind when it comes to analysis. This blank slate approach results in more novel findings that are firmly based on the participants experiences rather than being diluted by a priori assumptions. CGT also emphasises the importance of capturing the lived experiences of the participants. Given the lack of research incorporating the first-hand experiences of MbI, it is of paramount importance that the participants experience of MbI is captured and reflected in the findings. According to Charmaz (2006) the resonance of the findings for the participants is at the core the of CGT method, making it suited to the current study. The three stages of constructivist grounded theory analysis described by Charmaz (1990, 2006) were applied to the data. These are open coding, focused coding, and theoretical coding. The application of these coding practices to the data is described below. A sample of the coding process for the category 'Affection' within the core category 'Motivation' is also provided in Table 3.1 (See Appendix B for additional samples [2]). Data was continuously collected throughout the analysis stage until the saturation point was met whereby new data was not offering further insight into respondent's experiences of MbI.

The data was prepared for open coding by grouping the threads according to the online community they belonged. Line-by-line coding was then conducted whereby each line of data was given a label that captured the significant actions and processes. Charmaz (2006, p.50) guidelines for line-by-line coding were used to direct the labelling of each line of data. These guidelines included dividing the data into its constituent parts, defining the actions (using gerund verbs), uprooting implicit assumptions, deciphering inherent actions and meanings, clarifying the significance of the points, constant

[2] Verbatim text has been removed from Appendix B of the public version of the thesis to protect the privacy of the online community users.

comparison of data with data and identifying any gaps. Charmaz (2006) argues that one of the primary benefits of line-by-line coding is providing new leads that can be followed up and help to improve data collection. This was the case for the current study, as line-by-line coding helped to quickly reveal that MbI was not a term being used by the participants to refer to those faking illness online. Instead a variety of terms were being used to refer to the behaviour, for example, 'troll', 'faker', 'catfish', 'imposter' etc. Identifying the terms being used by participants meant that searches for relevant data became more efficient. Furthermore, line-by-line coding helped to reveal other communities where there had been cases of MbI and data from these communities was collected. Another function of line-by-line coding described by Charmaz (2006) is to correct any preconceived notions the researcher may have about the topic under investigation. This was facilitated by the constant comparison method and maintaining an open mind when coding and comparing data across all the communities. Using this method some early surprising directions of the research became apparent including the sophisticated methods and tenacious effort that was put in by the participants in identifying cases of MbI as well as the resiliency of the communities and the sympathy that some participants had towards those who perpetrated MbI.

Once the comparison of data across the communities was no longer yielding any new directions for research, focused coding began. Focused coding was conducted in two rounds. During the first round focused coding was conducted on the individual codes within the communities. The codes were organised according to frequency in combination with their theoretical significance. Round one of focused coding produced narrow categories, such as, 'giving the benefit of the doubt', 'risking false positives by confronting', 'presence of fakers disheartening', 'fakers stealing pictures and profiles', 'difficulty identifying fakers', 'falsely accused of being a faker' etc. Round two of focused coding concentrated on reducing the number of unique categories to create core categories. Firstly, the categories obtained from the individual communities were cross referenced so that similar categories could be grouped together. This often meant reformulating the categories obtained from the first round. For example, the discrete categories named variously as 'preferring to be duped than ignore genuine', 'others benefitting from advice regardless' and 'preferring to be trusting' all shared the common perspective that it was better to give the benefit of the

doubt when interacting with other members. Therefore, these discrete categories were amalgamated into the category ‘benefit of the doubt’. Secondly, the new amalgamated categories were cross referenced with one another to create core categories. This involved moving back and forth between the categories to identify ang group those that were theoretically similar on a more macro level. For example, ‘faker being punished through karma’, ‘fakers being pathetic people’, ‘faking being selfish behaviour’ and ‘faking is sick in a derogatory sense’, ‘describing faking as despicable’ were all included under the core category ‘unsympathetic’.

After the core categories were developed, theoretical coding began. According to Glaser (1978, p.72), the purpose of theoretical coding is to ‘weave the fractured story back together’. By weaving the story together, Charmaz (2006) describes how the analytical story emerges. This analytical story is developed by investigating the relationships between the categories. The core categories were cross referenced to establish connections that could be developed into an analytical narrative about the characteristics members had experienced through their encounters with MbI. For example, the analytical story related to the effect that MbI has on the community was brought together by combining the core categories ‘community as a support’ and ‘air of suspicion’, ‘witch hunts’ and ‘resilience’. This theoretical category produced two hypotheses relating to the effect that MbI has on an online community. It is hypothesised that MbI can either bring a community together or fracture it to the extent that the benefits of participation are eroded. Through constant comparison the overlapping relationships between the core categories were established to create theoretical categories. Based on these theoretical categories several hypotheses about MbI based on members encounters were developed. These hypotheses are discussed in-depth in the discussion section, as well as their relationship to existing literature.

Table 3.1 Sample of the Coding process for the theoretical category 'Effect of MbI on the community'

Initial Coding	Focus Coding Round One (Category)	Focus Coding Round Two (Core Category)
Kindness online extending offline Offerings of support when in a dark place Being online is their support system	Community offering an important source of support	Community offering the same characteristics as a support group.
Group wants all loss mothers to feel included Community is home from home Describing the community as a family	Close bond between the participants of the community	
Sharing things online they would not share offline Disclosing more online due to non-judgmental atmosphere Community as a place for talking about issues safely	Participants feel they can be more honest offline than online	
Suspecting participant of lying Questioning which other members are fake	Participants feeling suspicious of others and losing trust	The presence of perceived presence of MbI eroding the community by creating

Sad to lose faith in friendships in community		an air of suspicion.
Prevents them confiding in others	Discouraged from sharing experiences in the presence of fakers	
Distressed they may be sharing traumatic stories with liar		
Uneasy that fakers have access to personal stories		
Danger someone who needs help could be disbelieved	Less support offered out of the fear of being duped by a faker	
Fakers result in mistrust and participants not offering support		
Fears others in need may not receive support		
Moderator questioning their unusual circumstances	Falsely accused of being a fake	Witch hunts within the community results in participants restricting what they share to avoid being falsely accused.
False accusation was upsetting experience		
Told they were sad, pathetic and should kill themselves		
If they posted their family saga would be falsely accused	Censoring their experiences out of fear of being falsely accused of being a fake	
Do not share everything as it sounds like a soap opera		
If they shared everything others wouldn't believe them		
Hoping the risk of fakes does not deter support seeking	May discourage others seeking help due to fear of being falsely accused of being a fake	
Possibility of fakers discourages sympathy		
Doesn't want anyone to be dissuaded from seeking help		
Risks alienating genuine participants for sake of a few fakers	Discouraging other participants from going on 'witch hunts' that ultimately fragment	
Uneducated finger pointers are problematic		

Those who shout troll are no better than trolls Rather fall for a fake then not offer advice to a genuine Despite risk should still offer support and compassion Rather be duped than ignore someone in need Majority of people are honest on forum Majority are decent, trustworthy, and honorable	the community Gives benefit of the doubt and risk being duped rather than ignore someone in genuine need Believes despite the risk, majority are genuine	Community participants are resilient and choose to believe that the majority are genuine and even if duped it is a better alternative than ignoring a person in need of support.

Throughout the coding process analytical memos were written. Charmaz (2008) advocates the use of memos because they are the stepping stone between data collection and writing the research paper. In this way memos also offer a form of quality assurance by showing the thought processes that led from raw data to theory generation. Memos should provide an audit trail of the categories, contrasts, comparisons, questions, and avenues for future research that were discovered while analysing the data (Lempert, 2007). They offer the researcher an opportunity to interrogated the analysis process by offering a space to explore, investigate and develop ideas (Charmaz, 2006). Given the centrality of memo writing in CGT, analytic memos were kept during the coding process to provide a check for the decisions processes and an opportunity for reflexivity. A sample of the memos written is displayed below in Table 3.2 and 3.3 (See Appendix D for additional samples). Memos written during open coding explored data collection processes, potential categories, and theoretical directions. The memos written during focused and theoretical coding explored the development of core and theoretical categories.

Table 3.2 Sample of memos for open coding

Sample memos
MbI or Troll? It is important to distinguish between trolls and cases of MbI when collecting data. There have been instances of members referring to 'liars' who appear to have feigned illnesses or dramatic life events. However, these do not appear to be genuine cases of MbI as these liars tend to immediately start off with spurious stories and ask community members ridiculous questions, particularly around pregnancy. Some members speculate that these may be teenagers having a laugh and should be ignored. These types of liars appear to lose interest very quickly and disappear just as quickly as they arrived. Therefore, it is important to only include cases that closely adhere to the MO of a genuine MbI case. Ideally cases should only be included where the person has admitted to the behaviour, has been banned by a moderator (insider information) or have been caught red handed e.g. stole pictures, conflicting facts etc. *Spurious cues* In two communities, so far, there has been debate as to what cues are indicative of MbI. Some members claim any stories that are out of the ordinary or contain conflicting information are likely to be cases of MbI. However, another cohort of members strongly contests these

criteria. Particularly those who have been falsely accused of being an MbI perpetrator. They argue that experiencing dramatic life event is not aberrant and that many people have, what appears to others, extraordinary lives. This cohort also argues that providing conflicting information could be just a lapse in memory, particularly when it came to remember dates and times. People can't be expected to keep track of every detail. More concerningly is the indication that some members were withholding parts of their experiences that may appear extraordinary.in case they were accused of lying. This is a shame as the online community should offers a supportive space for members to be completely honest with one without the fear of being judged.

Table 3.3 Sample of memos written during focused and theoretical coding

Sample memos
Attitude towards MbI perpetrators The category 'attitude towards MbI perpetrators' is too narrow and needs to be broadened to offer some theoretical value, at present it is very descriptive. A possible way of dealing with this category is to connect attitude with perceived motivation. Attitude seems to be inextricably linked with perceived motivation. Those who are sympathetic and those who are unsympathetic attribute the behaviour to different causes. Both however tend to agree the goal of MbI is attention and sympathy (this insight into motivation matches the literature which may raise a red flag as to whether some of those commenting have personal experience of MbI). Sympathetic users view the behaviour as being caused by a mental illness. Unsympathetic users view the behaviour as being inexcusable and that these people are essentially just 'bad'. Combining attitude with motivation is analytically richer rather than keeping them apart where they are too descriptive and contribute little. *Cues of MbI* This is a large category that needs to be tidied up to uncover the nuances associated with cues of MbI. Cues of MbI contains several categories including 'gut feeling', 'impossible to imitate emotions', 'farfetched stories', 'incorrect details', 'stolen pictures', 'not existing offline' etc (titles of categories need to be fixed up). These need to be grouped together to create an analytical narrative. At present the best option seems to be to group them according to their evidential value. In other words, some cues are just indicative of MbI and can be regarded as circumstantial evidence, while other cues provide direct evidence that this is a

case of MbI. Within these two groupings there are cues associated with genuine members. There are two options for dealing with these, either create a separate category for them or use them as ‘opposite’ examples that can support the two groupings i.e. if meeting someone in person is a cue of genuiness than equally not being able to make contact with someone in the ‘real world’ is a cue of MbI.

3.4 RESULTS

Based on the analysis five theoretical categories and their respective core categories were created. These included attitude towards MbI based on perceived motivation (attention and sympathy, material exploitation, maliciousness, mental illness), personal effect of MbI (emotionally exploited and manipulated, unperturbed) effect of MbI on the community (community as a support group, air of suspicion, witch hunts, resilience) detecting MbI (types of MbI, clues which raise suspicion, clues associated with genuine users, clues which confirm suspicions) and managing MbI (difficult for moderators to control, publicly outing MbI users, report and ignore, prevention through caution and awareness).

3.4.1 PERCEIVED MOTIVATION

3.4.1.1 ATTENTION AND SYMPATHY

Of the users who described the types of scenarios that were faked by people with MbI, the most commonly reported, at 47%, was being pregnant or having an ill child. The fake pregnancies were typically complicated, involving family drama, multiples, miscarriages, and pre-mature births. Regarding having an ill child, this appeared to be a form of online by proxy, where the child had a chronic or unusual illness that eventually led to their death. The joint second most common scenarios faked, at 15%, were the MbI users own death which was typically a suicide and posting about a chronic illness, in most cases this was cancer. These two scenarios often involved the use of sockpuppets who would announce the death or post on behalf of the MbI user when they were hospitalised. Of the users who discussed motivation, 20% could not fathom why anyone would choose to 'experience' these undesirable life events. However, based on the emotionally eliciting scenarios feigned by MbI users, 17% speculated that the primary motivation was to receive attention and sympathy.

'Things I've discovered on the internet: Some people who are lonely and crave attention, want sympathy, spotlight, and the focus on themselves. If there is competition they are livid and try to outplay my illness in a "my illness is worse than yours" kind of way'

3.4.1.2 MALICIOUSNESS

32% of users were unsympathetic and took a hard-line position as they believed MbI was motivated by malice. They were disgusted and angered by the behaviour. MbI users were described as 'sick' in a derogatory sense and pathetic sad people who had nothing better to do with their lives. They viewed the behaviour as selfish because they fulfilled their own needs at the expense of those who were duped and detracted attention away from those who genuinely need help.

'This makes me so angry, there are people who truly reach out to the community because they need comfort and help from people in a similar situation, such as myself! I am sick that someone used this to get off, what a shame, I do not feel any bit of sympathy for this person, and all I have to say is: WHAT GOES AROUND COMES AROUND!'

3.4.1.3 MENTAL ILLNESS

In contrast 26% of users were sympathetic towards MbI as they believed the behaviour was precipitated by a mental illness which may be related to loneliness. These users felt that a person would have to be mentally ill to behave in this way and to invest so much energy and time into a fake persona. They felt sympathetic towards them as they were obviously experiencing some difficulties and they hoped they would receive the help they needed. Some users speculated about the root cause of the behaviour, believing that the MbI users may be lonely and were using fake personae to make friends online as they felt being themselves was insufficient.

'I hope that those who feel they need sympathy and attention based on false stories are able to find their way to getting the help they so badly need. I think those that fake a loss or an illness are so desperate to belong to a "group" that they will do anything.'

3.4.1.4 MATERIAL EXPLOITATION

4% of users reported materially exploitation, however it was not always the primary motivation. Users had spontaneously sent gifts and money to help them out during their time of peril, but in several instances MbI users used the illness or death of a child to solicit money for fictitious funerals, memorials, and medical expenses.

> *'A mom made up a story about her baby dying and some people started charity auctions for her and the less soft-hearted just bought stuff! They raised THOUSANDS. And then the mom went on a cruise with the money after she posted pictures of her dead baby. After the money had been sent the REAL mother of the dead baby came forward and was not happy that her child's image had been used in this matter. Gross.'*

3.4.2 PERSONAL EFFECT OF MBI

3.4.2.1 EMOTIONALLY EXPLOITED AND MANIPULATED

20% of users described feeling emotionally exploited and manipulated by MbI, with one member succinctly describing those with MbI as 'emotional vampires'. They had been emotionally invested in the faker to the extent that they grieved over fictitious deaths and had emotionally supported and consoled the MbI user throughout their various issues, genuinely caring about their welfare. In some instances, they were also going through the same issues as the MbI user and felt they had made an emotional ally. Once the ruse was exposed, those duped described the various emotions they experienced. These included humiliation, betrayal, violation, and anger. It was particularly distressing for those who had believed the faker was an ally in their shared experience.

> *'As her "illness" worsened, I drew her a beautiful card and mailed it the same day I heard the news from her sister that she had died. For days I was inconsolable.... it was not like me to emotionally invest in a person I have not known for long. But she really got me, hook, line, and ****** sinker. I will try to put this behind me and move on. She has taken enough from me already and from you all too.'*

3.4.2.2 UNPERTURBED

Another group, accounting for 8% of users, reported feeling unperturbed by MbI, viewing it as harmless. They could not understand why some users were becoming very upset over the behaviour as it was simply amusing to watch the scenarios unfold. However, this was an unpopular opinion to hold and those with this outlook were strongly criticised by other users who did not view it as benign.

> *'It's amusing that many of you say "Oh does it even matter? Just leave them be, they aren't hurting anyone!" ehhh...no? If you honestly believe that it is acceptable for people to lie in this way then you should sort out your priorities.'*

3.4.3 EFFECT OF MBI ON COMMUNITY

3.4.3.1 COMMUNITY AS A SUPPORT GROUP

14% of users described the importance of their online community as a support group. It served as an invaluable source of support and information for those with shared experiences. Some users felt they could be more open and honest online than offline when it came to expressing issues they were experiencing. Users also talked about the close bond they had with others, referring to the community as a 'family'.

> *'I came to this community looking for love and support from people who wouldn't look at me like I was odd for being me.... that's what I found here.... for this I am honestly grateful to you all! I have made some wonderful friends and found out that I am not alone in my journey.'*

3.4.3.2 AIR OF SUSPICION

Given the importance of the online communities as a source of support, the creation of fake personae was viewed negatively. 26% of users described how the presence or possible presence of MbI created an air of suspicion and speculation. Users talked about their belief that there were several MbI users in the community but they did not directly confront or name them. These silent nagging suspicions that there were MbI users among the community discouraged others from fully participating. This was even more pronounced if they had already been duped. They became hesitant to share their experiences, confide in others, or offer support and advice. This loss of trust and

cynicism led some users to fear that the sense of community and camaraderie was being slowly eroded.

'I spend my time on here giving people advice and getting close to them, I genuinely care about them and try to help out…. then to find out that it was all made up is disheartening and makes me unsure of everyone, I don't want to waste any more time investing in other people's problems.'

3.4.3.3 WITCH HUNTS

9% of users discussed the issue of witch hunts within the online community. The majority did not report directly confronting those they were suspicious of. However, several users did report being falsely accused of being an MbI user and being bullied by fellow users and moderators. The discussion around MbI led to some users wondering if they had ever been suspected. Some were even willing to verify their identity. Other users were already very conscious of this possibility and discussed their fear of being falsely accused which led them to censoring or restricting what they shared so as not to meet the perceived criteria of a fake. For example, they would not post about any deep personal issues they were experiencing and dramatic or unusual life events were edited to appear more typical so as not to raise suspicion. The fear of witch hunts also had some users questioning if it discouraged genuine users from seeking support. Overall, witch hunts were strongly discouraged by users in the absence of strong evidence for fear of falsely accusing someone who is reaching out for help and already vulnerable. Those who accused others were viewed negatively with some users suggesting that these people enjoyed the 'chase'.

'Some people have dramatic lives compared to others. They have been through a lot stuff and it might seem outrageous. I never thought when I posted something personal on here that others might be saying, "Yeah, right." My start in life was rough and because of that, my personal life all sorts of crazy, unique, sad, super cool, all before I was 25. Because of my own life experiences, I know life can sometimes be bizarre. When you're young with no family support, you can end up in one crazy experience after another.'

'I have been falsely accused of MbI, the moderator threatened to ban me from the community unless I scanned the receipt from my session at the neurologist and give the phone number so the moderator could call and confirm my story. I refused and they threatened to contact my facebook friends telling them I was lying about that person on an online community. It was the nasty and violent.'

3.4.3.4 RESILIENCE

While the presence or possible presence of MbI negatively impacted the community by creating an air of suspicion and an environment ripe for witch hunts, the communities were very resilient, with 22% of all users discussing this feature. They offered sympathy to those who had been duped, however, they felt that most users were genuine and they preferred to give others the benefit of the doubt and risk being duped rather than ignore someone who was genuinely in need of support. Some also felt that even if the person was an MbI user, someone else in need might read the advice they gave and it could be relevant to them.

'you just can't just let yourself get so paranoid that witch hunts are launched witch all over the place. I hate getting fooled as much as the next person, but I'd rather help someone on the chance that they need it, rather than being so suspicious that I let some real person slip through cracks.'

3.4.4 DETECTING MBI

3.4.4.1 TYPES OF MBI

Users described witnessing two types of MbI, of those 71% were covert and 29% overt. Covert is where the MbI users appear credible and convincing over a long period of time as there are few clues of deception. The users acknowledged the investment and commitment this took. When exposed, users of the online community are shocked as they were never under suspicion. In contrast, overt cases of MbI have very obvious clues of deception from the start, with wild inconsistency and implausible stories. Some users described overt MbI as simply amusing and often they would toy with them.

'I was shocked to find out it was all fake, at least the other ones were obvious fakes. Remember the one who went from posting "Do you think I might be pregnant?" to "I think I'm in labour!" to "I had my baby" in the space of an hour'

3.4.4.2 CLUES WHICH RAISE SUSPICION

29% of users described many clues which would make them suspicious. For 22% of those it was down to pure intuition. They simply felt that something was 'off'. For example, the MbI user would not be able to convincingly mimic the emotions associated with an issue as they had not lived it. This was sometimes sensed by users who had genuinely experienced the issue.

'we share such a life changing experience that you can never truly understand it until you have been through it and have paid the dues it costs. Therefore, fakers will always tend to betray themselves overtime.'

The other clues were more blatant, 34% pointed to overly dramatic and a-typical stories where there was a constant stream of negative and far-fetched stories.

'the pattern is the same, a very young, vulnerable single mother, who has issues with her own parents, their partner doesn't was to know, there are lots of problems during the pregnancy, reeling from one drama to the next, pregnancy with multiple babies, they go into labour ludicrously early, babies end up in special care, with one who didn't make it, constant updates about the ongoing health problems of the babies. If someone is telling you a story that sounds like it's from a Hollywood movie, then it probably isn't true'

33% believed inconsistencies and contradictions were indicative of MbI. The information conflicted with what they had previously said or was factually implausible. For example, the time-line given is unrealistic or their actions/behaviour fails to match that of someone in their position.

'what makes me very suspicious is that they pepper their posts with jargon you'd expect from a medical book or website instead of the usual words you'd expect from a person who is struggling, when someone inquires as to what the word means,

they either won't explain or give an inadequate answer. Their child has many other health issues and some are life-threatening but nothing in the photos seems to indicate that their child suffers from any of these issues.'

However, 14% of all users argued that dramatic/atypical stories and contradictions/inconsistencies were not necessarily indicative of MbI. They felt that under these criteria they could be accused of being an MbI user due to their dramatic/ atypical issues and the innocent contradictions/inconsistencies in their posts. In fact, some purposely did not reveal all the true details of their lives so they would not raise suspicion by appearing atypical.

'I have a life that I choose not to tell anyone about online because it sounds like a very bad soap opera and it's still ongoing. Once I am close friends with someone in real life, I will slowly let things slip. I think this colors my view that there are things in life that you have little to no control over and they can happen in weird ways.'

'sometimes I forget dates/times etc, that don't add up sometimes. I have had multiple miscarriages from 18 years ago up to more recently. The other day I posted and said it was 17 years ago, if someone was keeping track it would have seemed like I didn't get my facts straight and was lying! But I wasn't, I just forget the dates, it is possible for people telling the truth to get their facts confused.'

11% of those who discussed clues described another important clue, a friend or relative, who posts on behalf of the faker when they are pretending to be incapacitated due to illness or indeed have died. They believed that these were often sockpuppets. Again, some users argued that this was not necessarily unique to MbI users and a genuine user could ask a relative/friend to post on their behalf.

'I don't find it suspicious when family members post about an emergency or ask for prayers, but perhaps if it happened frequently I would. They wouldn't even need to put in a password because most forums save the login information so that you don't need to enter it each time you visit the site.'

3.4.4.3 CLUES ASSOCIATED WITH GENUINE USERS

Given that it was difficult to identify MbI based on clues of deception, 8% of users discussed if there were clues that would indicate if someone was genuine. These included meeting in real life and other users vouching for the member. However, it was argued that even this was fallible. There had been cases of people who enacted MbI online also continuing the ruse offline and those vouching for the user's credibility may in fact be sock puppets.

> *'This woman was convincing, she'd met about 300 people at the cancer meet up so it wasn't just online, people were concerned and upset for her. People who lost children or other family or suffered from cancer themselves were affected the worst. Because she played them and used their vulnerabilities by asking them for advice, not just online but one to one too.'*

> *'how many 'posters' are sockpuppets of the original troll? The more 'posters' you have endorsing the story, saying that they know the troll in real life or that they have had a similar experience, or that they are concerned about the troll, the more plausible the story appears. I reckon that's a far more frequent occurrence than people genuinely suddenly worrying for the troll.'*

3.4.4.4 CLUES WHICH CONFIRM SUSPICIONS

Once suspicions were raised, more solid evidence might naturally surface or be revealed through investigations conducted by users and/or moderators to confirm the suspicions. Unfortunately, this often didn't occur until a long-time into the deception. 9% of all users discussed these clues. Of these 42% described how some MbI users were found out as they had stolen photos and/or the identities of others who had online profiles.

> *'some person copied all my online photos and profile info and created a name like mine and then pretended to be me! They did it for years, occasionally I would find these profiles on sites like Bebo and Faceparty! According to this person, I was pregnant, miscarried, was abused by my parents, beaten around by my partner, I had a horrendous life LOL!'*

27% found that MbI users simply did not exist offline. There was no contact or records of their deaths or stays in hospital.

> *'The perpetrator was caught out through their sham funeral/death. There was no obituary and flowers could not be sent because there was no funeral home! Another simple way of discovering liars is to look for contact. Those posting crap avoid phone calls and visits because their lies cannot be supported and are very difficult to keep track of in person or in real-time.'*

31% noted that repeat offenders or those using sock puppets were easily identifiable because they shared a similar writing style and had the same IP address as the original.

> *'I could just tell it was the same member that was banned because of her by the writing style, attitude, and the topics she brought up, she was using multiple IP addresses but after running them, I knew for definitive it was her.'*

3.4.5 MANAGING MBI

3.4.5.1 DIFFICULT FOR MODERATORS TO CONTROL

41% of users discussed issues around managing MbI within their communities. Although of those 16% admired and praised the work that the moderators do in attempting to protect the community from MbI users, 14% believed that ultimately, they had little control.

> *'Thank you to the moderators for being so technically savvy in uncovering these trolls and for giving us all the facts. I am glad you have rules and are trying to keep us as safe as possible! Thank you moderators!'*

Moderators cannot vet and verify the identity of users and they cannot patrol the community trying to identify possible MbI cases 24/7. Furthermore, even if caught and banned, MbI users can simply re-join another online community or the same community by changing or disguising their IP address. There are no serious sanctions which can discourage the behaviour.

‘*Unfortunately, you will get these trolls on any forum, just report it and ignore it. We have a busy forum and both running and being a forum moderator is difficult work and most moderators have jobs and families etc.’*

‘I have no doubt that these kinds of trolls are in most online forums. There are no proper repercussions for them, so if they get banned they simply go on to the next forum.’

3.4.5.2 PUBLICLY OUTING MBI USERS

7% were pushing for moderators to publicly announce when there had been an MbI user on the forum rather than just banning and deleting the account without an explanation. In instances where moderators had publicly announced an incident in the community, the users expressed their gratitude. Users believed that there were several advantages to exposing cases of MbI, including helping people to recognise the common cues associated with MbI, raising awareness if the person is a repeat offender, shaming the MbI users and letting those who became emotionally invested know what had happened.

‘Great idea! If everyone knows who the troll is it'll completely take away the perverse excitement they get, once everyone finds out it was total bullshit. It'll also help others learn to recognize the king of threads trolls create as there are common red-flags to look out for in the posts. Great idea!’

A few users and moderators expressed their reluctance to publicly out MbI users. They felt that exposing them only fuelled their need for attention. Some cases were sensitive, e.g., minors and finally exposing them implies that the moderators are guaranteeing the genuineness of the remaining users, which is not the case.

‘just because a person is a member of this community, doesn't mean that they are trustworthy. Exposing the reasons for the ban would not really protect anyone, and in fact may create a misplaced sense of security by suggesting that we can somehow guarantee our members are genuine.’

3.4.5.3 REPORT AND IGNORE

In cases where users had directly confronted people they suspected of MbI, the outcome was mixed. They left the community, admitted to the behaviour and apologised or they became abusive. 25% believed that because there was a risk of falsely accusing a genuine member and the bullying that coincided with confrontation, it was best to report the behaviour to the moderators and ignore it. This allowed the moderators to investigate the situation and deal with it as they felt best.

'If people want to think someone is a troll that's fine, we are all allowed our opinion, but to continue bombarding them with private messages is too much, that's akin to bullying. if you suspect someone is a troll report them or ignore them because if it turns out they are not troll, you've destroyed what should been a great experience, finding out you're going to be a parent and sharing it. We should try to keep an open mind and be nice, if you suspect a troll report it and ignore it instead of starting a slagging match.'

3.4.5.4 PREVENTION THROUGH CAUTION AND AWARENESS

21% of users believed that MbI was a relatively common occurrence. Given that there was no fool proof way of ensuring that a person is genuine, 38% of users felt the best approach was to be cautious and aware. Being cautious involved having a healthy scepticism towards those met online, not becoming overly emotionally invested and maintaining boundaries.

> *'it should be "buyer beware" when deciding to become emotionally invested, online or offline. Don't get me wrong, these people are clearly disturbed, but doesn't every individual, and the community, have a responsibility to protect itself. Could it be that by entering a city environment, locking your door is just good sense?'*

Another facet of prevention was raising awareness about the possibility of MbI users to educate new users. However, some argued that this only resulted in scaremongering and paranoia which would deter participation.

> *'Being honest I'm not certain that this thread is helping anyone. New people are joining the forum all the time and the first thing they read is people faking illnesses*

etc. Who's to say what is a genuine case? I think most can tell the difference between a genuine case and a "tall story" Instead of giving the impression that this is rampant, we should just think what a sad incident and get on with trying to help each other.'

3.5 DISCUSSION

3.5.1 PORTRAYAL OF THE 'IDEAL VICTIM'

The users reported that the most commonly feigned scenario by an MbI user was a parent with a sick child or an expectant mother experiencing a difficult pregnancy, both targeting parenting forums. MbI users are falsely portraying themselves as what Christie (1986) refers to as an ideal victim, those 'who most readily are given the complete and legitimate status of being a victim' (pp18). They choose an ideal victim persona to maximise the benefits associated with occupying the sick/victim/hero roles, receiving sympathy, attention and external incentives. However, it's important to note that although external incentives were not generally the primary motivation for the behaviour, often users would offer to send money and gifts making it a consequence rather than a motivation, a distinction made by Worley et al. (2009). Ideal victims are typically women, children and the elderly and their images are often used in the media to elicit compassion from the audience (Hoijer, 2004). MbI users appear to be using a similar strategy by occupying personae that are viewed as ideal victims, mothers of sick children. The match between the persona and the target community is also important because it exploits three determinants of sympathy proposed by Loewenstein and Small (2007), (1) own personal state, users can easily relate to the parent/expectant parent persona as they share in the identity (2); past and vicarious experience, users are likely to have experience or knowledge of the issues feigned by being users of a parenting group, (3); in-group bias, strong sense of connectedness between users over shared situations and identities. Playing on these emotions ultimately activates an empathy response by mimicking the users in the community. Once the empathetic responses are aroused, not only will MbI users receive more sympathy and attention, their validity will be less likely to be questioned ultimately aiding in concealing their deceptions (Chen, 2012).

3.5.2 UNEXPECTED SYMPATHY TOWARDS MBI USERS

Users who were duped described feeling manipulated and exploited. This gave way to a variety of emotions including humiliation, betrayal, anger, and violation. The deception was particularity salient for those who were experiencing the issue being feigned by the MbI user. Feldman (2000) also reported on the anger and sadness felt by the users which in turn led to them trying to find the MbI user in the real world so they could vent their feelings and/or seek revenge, while others were simply unperturbed. This was not the predominant attitude found in this study as a significant proportion of users were sympathetic towards MbI users. Those who were sympathetic described feeling compassion towards those with MbI because they viewed them as mentally ill. Those who were unsympathetic were disgusted and angered by the behaviour and viewed the MbI users as selfish and 'sick' in a derogatory sense. The distinguishing feature between the two attitudes is whether the behaviour is attributed to mental illness. Therefore, the disparate reactions are consistent with Weiner's (1995) model of causal attribution. Corrigan et al. (2003) have successfully applied it to mental illness stigma. If a person is believed to be responsible for a negative event ("that person caused their own crazy behaviour") the reaction is one of anger. If a person is not believed to be responsible for a negative event ("He can't help it, he's mentally ill") the reaction is one of pity.

Despite the perceived origin of the behaviour explaining the disparate reactions, given the stigma surrounding FD, it is surprising that any users were sympathetic towards MbI users. Feldman (2004) has highlighted that FD is one of the most stigmatised mental disorders both within and outside the health profession, with no advocacy groups or funding for research into the disorder. Most likely because the behaviour associated with the disorder is a major social taboo. Those with FD are constantly violating implicit social contracts of trust associated with the sick role which can lead to frustration. For example, friends and family experience a range of negative emotions directed towards the FD sufferer including denial, anger, rage, disgust, and betrayal (Perlmutter, 2004). So, the question remains that despite the predominantly negative attitude in society towards FD, users expressed sympathy towards MbI users. This may be explained by looking at the other stigmatising factors, beyond causal attribution, considering MbI being more remotely experienced. Direct contact with an

FD sufferer may magnify factors which contribute to a stigmatised attitude. These include peril, (behaviour associated with FD is more frightening, unpredictable and strange when FtF), visibility of the illness (online users can immediately cut contact, however family, friends and health professionals are invested and cope with the behaviour long-term), course and stability (the prognosis for FD recovery is poor, those closely acquainted with an FD sufferer are more likely to be aware of this than an online user who servers contact) and disruptiveness (FD has a significant negative effect on relationships and success in society, this would be of little consequence to online users) (Ahmedani, 2011). The increased exposure to the FD sufferer in comparison to the MbI user may help to account for the more stigmatised attitude towards FD in comparison to MbI.

3.5.3 SELF-CENSORSHIP LIMITING OPPORTUNITIES TO CONFER COMMUNAL NORMALITY

MbI not only impacts the individual users who encounter them but also the online community. The users viewed their online community more as a support group than a community. The presence or possible presence of MbI creates an air of suspicion amongst the users as to who can and cannot be trusted. This led to some users being discouraged from fully participating, sharing experiences confiding in others and offering and seeking advice. However, it did not have a deleterious effect and the communities were resilient. Resiliency was not attributable to solidifying and reinforcing the norms of the community through the pursuit of MbI users, as is the case with trolls (Herring et al., 2002; Kuntsman, 2007). But rather the users reported that the community norm of helping outweighed the risk of being duped. They described preferring to give the benefit of the doubt and risk being duped rather than ignore someone in need.

This drop off from participation would be expected given the importance of interpersonal trust in maintaining online communities. But the erosion of interpersonal trust is also harmful to the functioning of the online community as a support group. Interpersonal trust determines if a member will make the initial risky emotional investment to participate and it also helps to develop personal relationships that have supportive meaning and are beneficial to the user (Helgeson & Gottlieb, 2006; Klein

& Dinger, 2008). This study also found that not only does the erosion of interpersonal trust limit participation and its benefits, as would be predicted it also has a secondary unexpected effect. The erosion of interpersonal trust was further compounded if users became cognizant that they could be under suspicion themselves. There were instances of users, who claimed they were genuine, reportedly being bullied and harassed by other users who did not believe them. The fear of being falsely accused resulted in some users altering and editing their experiences so they would not overlap with the criteria associated with MbI i.e. dramatic and a-typical. A judgemental atmosphere in a support group is harmful because it prevents users freely sharing difficult emotions, experiences and issues which could be partially alleviated by other users confirming their normality (MacCarthy et al., 1989; Hall & Lloyd, 1993; Ahlberg & Nordner 2006). Those who are afraid of being falsely accused ensure that what they share conforms to the common narrative of the group. This means that valid emotions, issues and experiences that they perceive to be a-typical and/or dramatic remain suppressed despite the potential benefit it would have for themselves and other users by confirming normality. This suppression has a deleterious effect on one of most championed benefits of online communities which is facilitated by anonymity. Providing a safe realm for situations characterised by stigma, loneliness, silence, and health concerns to be freely discussed, in turn conferring communal normalisation (White & Dorman, 2001; Trondsen & Tjora, 2015). Ironically, the anonymity which facilitates this process also appears to hinder it when there are witch hunts for MbI users in an online community, preventing honest self-disclosures.

3.5.4 DIFFICULTY IDENTIFYING MBI USERS

Given the concerns around trust and not wanting to appear to be an MbI user, there were various suggestions as to what clues were indicative of MbI. The most frequently cited were dramatic/a-typical stories and contradictions/inconsistencies which were also noted by Feldman (2000). These are referred to as high-level clues within the deception literature. However, they are not definitive and can be inadvertently created by genuine users (Rowe, 2005). The users also recognised this issue by strongly arguing that these criteria could easily apply to themselves. Therefore, more solid clues were needed to establish with certainty that it was an MbI user. One of the most reliable methods of detecting deception is checks of authoritative references (Rowe,

2005). In fact, users and moderators were already actively engaged in checking authoritative references to confirm or allay their suspicions. This included checking IP addresses to see if it was shared among users to identify sock puppets, investigating if the person existed offline and searching the internet to see if photos or whole identities had been stolen. In some instances, these clues naturally surfaced overtime. Conversely, users also discussed cues of trustworthiness. These included meeting the person face-to-face or another user vouching for their credibility. Both however were viewed as fallible, people have been known to act out MbI in real life and those vouching for the user could be sock puppets.

The current approach to identifying MbI is unsatisfactory. The high-level clues are ambiguous and the checking of or surfacing of authoritative references takes time and occurs too late into the deception when the damage has already been done. Although users were appreciative of moderators, they felt there was little they could do to combat the behaviour. They could not vet all users or monitor the community 24/7 and even if banned, MbI users could return under a new IP address. A new approach needs to be adopted to help identify cases of MbI quicker and more reliably. One option that could be used to aid in the early detection of MbI is the use of automated methods of deception detection such as a text classifiers developed through supervised machine learning. They have already been used successfully to detect child exploiting chats, deceptive opinions in reviews, stylistic deception, gender deception, spam detection and malicious profiles (Miah et al., 2011; Ott et al., 2011; Afroz et al., 2012; Alowibdi et al., 2015; McCord & Chuah, 2011; Fire et al., 2012). A second option is a system built into the online community that requires all images to be reverse searched before posting. This would help to automatically detect users who are stealing other people's photos and/or online profiles and trying to pass them off as their own. These methods would relieve moderators and users from the burdensome task of trying to identify cases of MbI.

3.5.5 No deterrents for MbI

Once a suspected case of MbI has been identified, users and moderators have to decide on the best course of action. The majority believed the best policy was to report the user to the moderators and then ignore them. This is because in cases where the MbI

user had been directly confronted, the results were mixed, which was also reported by Feldman (2000). There was also the danger that a genuine user could be falsely accused, leading to fallout within the community. This sage advice is also commonly given when dealing with disruptive online users such as trolls (Wallace, 2015). Once moderators have investigated and confirmed a case of MbI some users believed that they should be publicly outed. There has already been a legal precedent set allowing users to expose the identities of MbI users in the public interest without fear of being sued for defamation, once deception has occurred (Feldman & Peychers, 2007). However, moderators were particularly reluctant to publicly out MbI users as some of the cases involved vulnerable people and to do so would be insensitive. This highlights a key issue as to whether MbI should be primarily treated as a cybercrime. Feldman and Peychers (2007) and Pulman and Taylor (2012) argue that it should be, by suggesting that it is legally pursued. Given that the actions of FD sufferers are rarely categorised as criminal and legally pursued, it is difficult to envisage how it's online equivalent could be treated as a crime at present (Miller, 2007). This is further confounded by the fact that legislation and regulation around cybercrime is still in its infancy. Currently there is no deterrent or punitive action that can be taken against MbI.

3.5.6. PREVENTATIVE ACTION

Given the difficulty involved in identifying and eliminating MbI, users advocated a more preventative approach by being more cautious. They believed in maintaining a healthy level of scepticism towards other users and maintaining boundaries by not becoming overly emotionally involved. This is a sensible approach given the current issues with managing MbI. However, this emotionally guarded stance towards other users may have a negative impact on interpersonal support available within the online community. The main features of which are attachment, obtaining guidance, social integration, reassurance of worth, a sense of reliable alliance and opportunity for nurturance (Weiss, 1974). User also suggested raising greater awareness about MbI, some moderators and users were opposed to this as they thought it would deter participation. Another form of preventative action could be in discouraging MbI users from joining the online community to begin with. A number of methods used to

discourage trolls could also be applicable to MbI users by deferring immediate gratification (Grohol, 2006; Hall, 2013; Kiesler, 2010):

- Require registration before posting to deter casual troublemakers.
- Use a verification system so that the email address must be validated.
- Publicly display the originating IP address on every post to increase accountability.
- Require a probationary period before a user can post.
- Require new users to post bonds that can be refunded if they are deemed to be reputable overtime.

While these are certainly not fool proof in safeguarding an online community, they may help to reduce vulnerability to such behaviour by making the membership process more onerous.

3.6 CONCLUSION

Identity deception in online communities erodes interpersonal trust which jeopardises the benefits of participation. One such form of identity deception that had received little research attention is MbI. The purpose of this study was to explore MbI from the perspective of a large sample of online community users. The findings highlight how MbI users purposely maximise the attention and sympathy they can receive and increase their validity by preying on the empathy of online community users by occupying 'ideal victim' personae. Despite the calculated nature of the behaviour a significant proportion of users were sympathetic towards MbI. This may be accounted for by their attribution of the behaviour to a mental illness and lack of physical proximity. The most worrying consequence of MbI was that not only does it erode interpersonal trust thus limiting full participation it also creates a fear of being falsely accused of MbI. Those in fear edit the experiences they share which limits opportunities to confer normality, an important benefit of online community participation. Unfortunately, users find this nuisance behaviour in online communities to not be easily identifiable and believe there is little that can be done to deter it. Given the negative impact of MbI on online communities, solutions to these issues need to be found to aid early detection and discourage the behaviour. Suggestions including a

text classifier, reverse photo search filter and a more onerous membership process have been made but future research needs to expand on these options and devise how they can be practically incorporated into online communities.

CHAPTER 4.0 DETECTING MUNCHAUSEN BY INTERNET: DEVELOPMENT OF A TEXT CLASSIFIER THROUGH MACHINE LEARNING.

Aideen Lawlor * & Dr Jurek Kirakowski

School of Applied Psychology,
University College Cork

Submitted for review to the *Journal of Communication*

Statement of contributions

Lawlor, A. (Candidate)
Writing and compilation of manuscript, established methodology, data collection, data analysis and data interpretation.

Kirakowski, J. (Supervisor)
Supervised development of work and editing of manuscript.

This chapter is a revised version of the journal paper referred to above

4.1 ABSTRACT

Munchausen by Internet (MbI) has a negative impact on online communities by emotionally exploiting members into caring for people who create fake stories. At present, there is no method of systematically identifying MbI. This study aimed to develop such a method, by applying the SLP (Social Language Processing) (Hancock et al., 2010) paradigm to the problem. This involved identifying and extracting discourse features and training a classifier through machine learning, to identify if text was written by a genuine person or a person exhibiting Munchausen's. A corpus of text written by people falsely and genuinely occupying the sick/victim/hero role was collected from the internet. Two models were created for training. The first model was based on significant Linguistic Inquiry Word Count 2015 (LIWC2015) dimensions and the second, a n-gram model obtained through vectorisation. The best prediction accuracy for both was obtained using an AttributeSelectedClassifier with Naïve Bayes. The LIWC2015 model had a prediction accuracy of 81.11% and the N-gram model, 81.67%. While the overall prediction accuracy is high, Munchausen's was incorrectly classified as Genuine in approximately 30% of cases. A larger corpus and inclusion of the dyadic relationship is needed to lessen the number of false negatives. Beyond the development of a viable classifier, the LIWC2015 predictive attributes obtained support the growing view within psychological theories of deception that domain specificity must be considered when interpreting discriminate cues of deception. This is because there is no one-size-fits-all model of LBC's of deception, they are context dependent. LBC's of deception specific to MbI relate to motivation and proficiency, creating and maintaining immediacy as well as the need to stress social processes. The discriminate cues obtained also support existing LBC's of deception theories specifically those pertaining to the reduction of cognitive complexity.

Key words: Munchausen by Internet; Online identity deception; Text classifier; Machine Learning

4.2 INTRODUCTION

Munchausen by Internet (MbI) is a form of online identity deception where a person falsely occupies the sick/victim/hero role through their online postings. Lawlor and Kirakowski (2016) noted that, MbI users tended to occupy 'ideal victim' personae defined by Christie (1986) as those 'who most readily are given the complete and legitimate status of being a victim' (pp18). The most common persona adopted by MbI users is that of a mother with a sick child or a mother experiencing a difficult pregnancy, often ending in a miscarriage. MbI users violate the trust which is automatically bestowed upon those who are 'ideal victims', causing emotional distress to those caught up in their web of deceit. In addition to the distress caused it also creates distrust. Lawlor and Kirakowski (2016) found that the presence or possible presence of MbI users created an air of suspicion in online communities. This led to more reserved participation to avoid being deceived or falsely accused of being an MbI user. More reserved participation maybe deleterious to the therapeutic benefits of online communities, by limiting the opportunities to confer normality and cultivate interpersonal support. MbI can also have serious health consequences for those who follow the ill-informed advice or misinformation provided by MbI users (Witney, Hendricks & Cope, 2015). In some instances, MbI can be primarily motivated by monetary gain or simply be a consequence, leading to other users being materially and monetarily exploited (Lawlor & Kirakowski 2016; Worley, Feldman & Hamilton, 2009). Given the undesirable effect that MbI users have on an online community, being able to reliably detect them is of paramount importance.

Lawlor and Kirakowski (2016) found that at present MbI users are typically suspected based on high-level deception cues and these suspicions are confirmed or discredited by checking authoritative references. The high-level deception cues are dramatic/a-typical stories and contradictions/inconsistencies. Authoritative references include checking IP addresses for duplicate users (sock puppets), looking for evidence that the person and their claims exist in the real world and checking if photos or whole identities have been stolen from another user. These authoritative references are ultimately the gold standard for identifying cases of MbI. The initial high-level deception cues are however problematic, as they can be inadvertently applied to those who are genuine. This was noted by the participants in Lawlor and Kirakowski's

(2016) study, where participants believed the cues of deception could be applicable to them and some even edited their experiences when posting to conform to the typical experiential narrative of the community.

Given the ambiguity of high-level deception cues, a more discerning method of identifying possible cases of MbI is needed to try and lessen the number of genuine users who fall under suspicion and the amount of MbI users who evade suspicion. This would increase the likelihood of a suspected user being an MbI user, which would help concentrate the checking of authoritative references. Rather than relying on human judgement, which can be biased, this study will employ computational linguistic techniques to try and build a classifier to predict whether text was written by a person exhibiting Munchausen or a genuine writer. Two models are tested, one using a classifier derived from LIWC2015 and one from the N-gram model.

4.2.1 Social Language Processing (SLP) paradigm

The stages involved in developing a classifier described by SLP (Hancock et al., 2010) were applied to the study. SLP offers an interdisciplinary (cognitive science, communications, computational linguistics, discourse processing, language studies and social psychology) approach of examining the features of communication using three recursive stages:

1. *Identifying potential language features* – Based on theory and existing research, identifying what the relevant cues are that underscore the social dynamic (deception).
2. *Automatically extracting the relevant discourse features* – Using a software programme to extract these relevant features from the document.
3. *Developing statistical classifiers* – Creating a classifier that can discriminate one type of document (genuine) from another (deceptive) based on the relevant discourse features selected.

4.2.2 Stage 1: Identifying potential language features

4.2.2.1 Deception theories

There are several deception theories that underpin LBC of deception, including

Criteria-Based Content Analysis (CBCA), Reality Monitoring (RM), Verbal Immediacy (VI) and Cognitive Load Hypothesis (CLH) which all feed into Interpersonal Deception Theory (IDT). CBCA is based on Undeutsch's (1989) hypothesis that a statement made from memory will be different in content and quality than one created from imagination. Steller and Köhnken (1989) went on to develop a checklist of 19 CBCA criteria that could be used evaluate the credibility of a witness statements based on Undeutsch's (1989) hypothesis. The 19 criteria are centred around five categories and are scored on a three-point scale by a trained investigator. A score of "0" is given if the criterion is absent, "1" if the criterion is present, and "2" if the criterion is strongly present. Table 1 provides a list of the 19 criterion which are centred around five categories.

Table 4.1 Criterion for CBCA

19 CBCA criterion
General Characteristics
1. Logical structure
2. Unstructured production
3. Quantity of details
Specific Contexts
4. Contextual embedding
5. Descriptions of interactions
6. Reproduction of conversation
7. Unexpected complications during incident
8. Unusual details
9. Superfluous details
10. Accurately reported details misunderstood

11. Related external associations

12. Accounts of subjective mental states

13. Attributions of perpetrators mental state

Motivation related contents

14. Spontaneous corrections

15. Admitting lack of memory

16. Raising doubts about one's own testimony

17. Self-deprecation

18. Pardoning the perpetrator

Offence specific elements

19. Details characteristic of the offence

In a meta-analysis by Amado, Arce and Fariña, 2015, the CBCA criteria were shown to be consistently successfully in discriminating between memories of self-experienced real-life events and fabricated or fictitious events. The criteria *quantity of details* and *details characteristic of the offence* had the highest discriminatory power. These were followed by the categories *logical structure*, *unstructured production*, *contextual embedding*, *description of interactions*, *reproduction of conversations*, and *accounts of subjective mental states* which all had medium discriminatory power. The remaining categories had low discriminatory power.

Johnson and Raye's (1981) theory of RM proposed that the source of a person's memory can be ascertained through the qualitative characteristics of their descriptions of an event. The two sources of memories are external memories (reality) and internal memories (fantasy). The qualitative characteristics that are used to discriminate between external and internal memories are contextual information (time and place), sensory information (colours, sounds, smells etc.), semantic information (such as story repetition) and cognitive operations (thoughts and reasonings). Johnson and Raye

(1981) hypothesised that because external memories are derived from perceptual processes, they will include more contextual, semantic, and sensory information than internally derived memories. In comparison, internal memories are derived from internal processes and therefore will include more cognitive operations. Unlike CBCA, RM does not have a formal list of criteria for assessing the veracity of statements. However, Sporer and Kupper (1995) used the RM to develop the Judgement of Memory Characteristics Questionnaire (JMCQ) that could be used by a trained rater to detect self-experienced and invented events. The JMCQ is made up of eight scales which are called the RM criteria. These include clarity/vividness, sensory experiences, spatial information, time information, emotions and feelings, reconstruct ability of the story, realism and cognitive operations. A review by Masip et al. (2005) found that the RM system discriminated reasonably between self-experienced and invented events at above chance levels. However, the individual discriminatory power of the criteria was mixed, with the most consistently discriminative criteria being visual and auditory details, contextual information, time information and realism. Cognitive operations were found to be associated with both truthful and deceptive accounts, however this inconsistency may be attributable to moderator variables associated with different designs of the studies e.g. different definitions of criteria.

The theory of VI was developed by Wiener and Mehrabian (1968) to detect the valence of affect i.e. whether a person likes or dislikes what they are talking about. The valence is discerned through the persons choice of words or phrases that indicate either approach or avoidance of the referents of discourse. Kuiken (1981) extended VI theory to deception detection. It was proposed that because deception is associated with negative affect and dissociation (distancing themselves from the lie), low levels of verbal immediacy would indicate that that a person is lying. Subsequent studies also found a consistent relationship between low levels of verbal immediacy and deception (DePaulo, 1981; Zuckerman, DePaulo & Rosenthal, 1981; Zuckerman & Driver, 1985; DePaulo et al., 2003; Buller et al., 1996). These studies typically draw from the criteria developed by Weiner and Mehrabian (1968) to operationalise high and low verbal immediacy messages. Five criteria underpin the scales used for scoring the level of verbal nonimmediacy. These are spacial and temporal separation, denotative specificity, selective emphasis, agent-action-object-relationships, and modifications. Spacial and temporal separation refers to the communicator distancing themselves

from the object of communication. For example, using introductory phrases such *"At the beginning I was writing"* instead of the more immediate *"I was writing"*. Denotative specificity is divided into two categories, over and under inclusion. Over inclusion is being non-specific. For example, to the question *"who took care of you?"* a person could say *"she always did"*. The word *'she'* could refer to any female and is therefore regarded as overly inclusive. In contrast, under inclusion is too specific. For example, a person could answer *"I don't like her manners"* to the question *"Do you like Jane"*. However, the symbol in the question is *'John'* and not *'manners'*. Therefore, there is a mismatch between the answer and the question due to the use of an under inclusive symbol in the answer. Selective emphasis refers to several variations including the sequence in which objects are introduced, the frequency, intensity, extensity ascribed to an event and over and under responsiveness. Examples of frequency, intensity and extensity modifiers include *'little'*, *'great'*, *'much'*, *'very'*, *'seldom'*, *'often'*, or *'never'*. The use of such modifiers reveals the communicators affect, evaluation or preference. Agent-action-object-relationships refers to the extent that the communicator is positively or negatively disposed towards an activity, person involved in the act or addressee. For example, saying *"I had to go"* rather than *"I went"* denotes a negative attitude towards leaving. Modification is expressed through qualification or objectification. Qualification indicates uncertainty by using phrases such as *"I think"*, *"it might be"* or *"it could be"*. In contrast, the use of objectification phrases such as *"it is obvious"*, *"it is true"* or *"it is obvious"* indicates certainty.

CLH assumes that telling a lie is more cognitively demanding than telling the truth (Zuckerman, DePaulo, & Rosenthal, 1981). This is because a lie is a novel creation, it requires effort to make sure its internal content is consistent and consistent with what the liar has previously said. This increases cognitive load resulting in the spill over of leakage cues in the form of reduced cognitive complexity. Vrij (2015) highlights the wide variety of criteria have been used across numerous studies to successfully detect the increase in cognitive demand associated with deception. Studies investigating the criteria for assessing CLH typically involve interventions that increase the cognitive load which is followed up by unanticipated questions. Under these conditions, truth tellers to liars are compared to identify the cues of deception associated with the CLH. Using this method numerous criteria have been identified including increased pauses, decreased blinking, decreased hand and finger movements, increased gaze aversion,

providing fewer details, less plausible answers, more contradictions, less overlap in statements given (Vrij & Mann, 2003; Kohnken, 2004; Leins, Fisher & Vrij, 2012). The linguistic cues of cognitive complexity identified include using longer sentences, syntactically complex sentences, and a decrease in exclusive words and increase in motion verbs (Depaulo et al., 2003; Toma & Hancock, 2012).

The final framework, IDT (Burgoon & Buller, 1996), pulls these preceding theories together and adds to them to create a comprehensive theory of deception. IDT proposes that liars try to strategically portray themselves as being truthful but ultimately will leak non-strategic cues of deception. However, a dynamic process unfolds whereby if the liar can tell that the receiver has become suspicious they can learn and revise their strategy to become more proficient. In liars' attempts to portray themselves as truthful they produce strategic cues of deception these include (Buller & Burgoon, 1994; Vartapetiance & Gillam, 2012; Zhou et al., 2004):

1. *Uncertainty and vagueness* – Creating ambiguity through short and non-committal language, so there is nothing concrete to be challenged. Linguistic cues include passive voice, indefinite pronouns, fewer self-references and fewer absolute verbs.
2. *Non immediacy* – Symbolically trying to remove oneself from the situation and create distance. Linguistic cues include changing tense from present to past, reference, modal verbs, uncertainty, objectification and generalising terms.
3. *Dissociation* – Trying to divert responsibility by distancing oneself from the deceptive act. Linguistic cues include increased use of levellers, group/other references, modifiers and fewer self-references.
4. *Image and relationship protection* – Deceivers try to disguise the cues that could expose their deception, negatively impacting their reputations and relationships, by trying to appear sincere. Linguistic cues include the avoidance of negative affect and discrediting information.

In addition to strategic cues, liars also produce non-strategic cues which leak out due to the high cognitive load endured as they attempt to avoid detection. For the most part, the non-strategic cues proposed by IDT are non-verbal. However, Zhou et al. (2004) highlight two exceptions which can be revealed through linguistic cues,

unpleasantness, and inexpressiveness. Both are evident by adverbs and adjectives that denote negative affect and the use of less expressive or intense language.

4.2.2.3 USING LIWC TO CREATE A LINGUISTIC PROFILE OF DECEPTION

The deception theories presented provide a method of interpreting linguistic cues of deception. However, the criteria used by these theories to detect deception are not always compatible with detecting online deception which is largely text based. Newman et linguistic profile of deceptive language that is based on deception theories. For example, Newman et al. (2003) developed a unique solution to this problem by using a computer based text analysis programme, which analyses the linguistic properties of a text, to create a that cognitive demand could be identified through fewer exclusive words (*but, without, except*) and more motion verbs (*walk, go, carry*). This is because the use of exclusive words involves making distinctions between what did and did not happen, a task that is too cognitively demanding for a liar. The inclusion of more motion verbs is an attempt to add credibility to a story that involves little cognitive demand, compared to more complex words associated with evaluations and judgements.

The LIWC developed by Pennebaker et al. (2015) has been used by numerous studies to develop linguistic profiles of deception that are informed by deception theories (see table 4.3). LIWC searches through the words or word stems of a given text and then categorises them into a language dimension (see table 4.2).

Table 4.2 LIWC Language dimension

Dimension	Example	Psychological Correlate
Word count		Talkativeness, verbal Fluency
Words per sentence		Verbal fluency, cognitive complexity
Dictionary words	Percentage of all words captured by the program	Informal, nontechnical language
Words < 6 letters	Percentage of all words longer than 6 letters	Education, social class

Total function words		
Total pronouns	I, them, itself	Informal, personal
Personal pronouns	I, them, her	Personal, social
First-person singular	I, me, mine	Honest, depressed, low status, personal, emotional, informal
First-person plural	We, us, our	Detached, high status, socially connected to group
Second person	You, your, thou	Social, elevated status
Third-person singular	She, her, him	Social interests, social support
Third-person plural	They, their, they'd	Social interests, out-group awareness
Indefinite pronouns	It, it's, those	
Articles	A, an, the	Use of concrete nouns, interest in objects and things
Common verbs	Walk, went, see	
Auxiliary verbs	Am, will, have	Informal, passive voice
Past tense	Went, ran, had	Focus on the past
Present tense	Is, does, hear	Living in the here and now
Future tense	Will, gonna	Future and goal oriented
Adverbs	Very, really, quickly	
Prepositions	To, with, above	Education, concern with precision
Conjunctions	And, but, whereas	
Negations	No, not, never	Inhibition
Quantifiers	Few, many, much	
Swear words	Damn, piss, fuck	Informal, aggression,

Psychological Processes		
Social processes	Mate, talk, they, child	Social concerns, social support
Family	Daughter, husband	
Friends	Buddy, friend, neighbour	
Humans	Adult, baby, boy	
Affective processes	Happy, cried, abandon	Emotionality
Positive emotion	Love, nice, sweet	
Negative emotion	Hurt, ugly, nasty	
Anxiety	Worried, nervous	
Anger	Hate, kill, annoyed	
Sadness	Crying, grief, sad	
Cognitive processes	Cause, know, ought	
Insight	Think, know, consider	
Causation	Because, effect, hence	
Discrepancy	Should, would, could	
Tentative	Maybe, perhaps, guess	
Certainty	Always, never	Social/verbal skills, emotional stability
Inhibition	Block, constrain, stop	
Inclusive	And, with, include	
Exclusive	But, without, exclude	Cognitive complexity, honesty
Perceptual processes	Observing, heard, feeling	
See	View, saw, seen	
Hear	Listen, hearing	
Feel	Feels, touch	
Biological processes	Eat, blood, pain	

Body	Cheek, hands, spit	
Health	Clinic, flu, pill	
Sexual	Horny, love, incest	
Ingestion	Dish, eat, pizza	
Relativity	Area, bend, go	
Motion	Arrive, car, go	
Space	Down, in, thin	
Time	End, until, season	
Personal concerns		
Work	Job, majors, Xerox	
Achievement	Earn, hero, win	
Leisure	Cook, chat, movie	
Home	Apartment, kitchen, family	
Money	Audit, cash, owe	
Religion	Altar, church, mosque	
Death	Bury, coffin, kill	
Spoken categories		
Assent	Agree, OK, yes	Agreement, passivity
Nonfluencies	Er, hm, umm	
Fillers	Blah, Imean, yaknow	Informal, Unprepared speech

Table 4.3 Discriminate LIWC deception cues obtained from research.

LIWC Cue	Rate	Theory/ Hypothesis	Data source	Authors
Self-references	Fewer	IDT, Strategic cues of dissociation, uncertainty/vagueness and nonimmediacy.	Online dating profiles	Toma & Hancock (2010)
			Meta-analysis	Hauch et al. (2015)
			Online dating profiles	Hancock et al. (2008)
			Opinions, feelings	Mihalcea & Strapparava, (2009)
			Attitudes, feelings, denials	Newman et al. (2003)
			Prisoners statements	Bond & Lee (2005)
			Music preferences	Dzindolet & Pierce (2004)
			Online dating profiles	Toma & Hancock (2012)
	More	Inserting themselves in a story to create credibility, lacking guilt for lying	Online opinion spam	Ott et al. (2011)
			High fantasy prone	Schelleman-Offermans & Merckelbach (2010)
			Enron emails	Gupta & Skillcorn (2006)
			Psychopaths	Almela, Alcaraz-Marmol & Cantos (2015)
			Online opinion spam	Li et al. (2014)
Other-references	Fewer	May be using specific names or nouns (sister, mother etc.) rather than vague pronouns to try and add credibility.	Meta-analysis	Hauch et al. (2012)
			Attitudes, feelings, denials	Newman (2003)
			Prisoners statements	Bond & Lee (2005)
			Music preferences	Dzindolet & Pierce (2015)
	More	IDT, Strategic cues of dissociation.	Online dating profiles	Hancock et al. (2008)
Sense words	Fewer	CBCA & RM, lack the real experience for perceptual details.	Meta-analysis	Hauch et al. (2015)

	More	Enhancing credibility by including extra detailed descriptions.	Online dating profiles	Hancock et al. (2008)
Exclusive words	Fewer	CLH, supplying complex information about what did or did not happen increases the cognitive load.	Meta-analysis Attitudes, feeling, denials Prisoner statements	Hauch et al. (2012) Newman (2003) Bond & Lee (2005)
	More	Proficient liars may intuitively understand that the use of exclusive words will impress as powerless speech.	High fantasy prone	Schelleman-Offermans & Merckelbach (2010)
Negation words	More	IDT, Dissociation from the lie through the use of non-committal language.	Meta-analysis Online dating profiles Online dating profiles Emergency calls	Hauch et al. (2012) Toma & Hancock (2010) Toma & Hancock (2008) Burns & Moffitt (2014)
Relativity words	Fewer	CBCA & RM, lack the real experience to describe. CLH creating these additional details would increase cognitive load.	Meta-analysis	Hauch et al. (2012)
Positive affect	More	IDT, image and relationship protection, trying to appear more sincere and confident to disguise deception.Over exaggerating feigned sentiment.	Meta-analysis Financially motivated Psychopaths Online opinion spam	Hauch et al. (2012) Hirschberg et al. (2005) Almela, Alcaraz-Marmol & Cantos (2015) Li et al. (2014)

Negative affect	Fewer	IDT, image and relationship protection, trying to appear more sincere to disguise deception.	Online dating profiles Online dating profiles	Toma & Hancock (2010) Toma & Hancock (2012)
Assent words	More	Limit and control answers.	Emergency calls	Burns & Moffitt (2014)
	More	May indicate guilt about lying and/or about the topic of their lie. Over exaggerating feigned sentiment.	Meta-analysis Attitudes, feeling, denials Enron emails Opinions, feeling Prisoner statements Music preferences Psychopaths Online opinion spam	Hauch et al. (2012) Newman et al. (2003) Gupta & Skillcorn (2006) Almela,Valencia-Garcia & Cantos (2012) Bond & Lee (2005) Dzindolet & Pierce (2004) Almela, Alcaraz-Marmol & Cantos (2015) Li et al. (2014)
Motion verbs	More	CLH, lower the cognitive complexity by using straightforward descriptions that avoid evaluation and judgement.	Meta-analysis Attitudes, feelings, denials Enron emails Prisoner statements	Hauch et al. (2012) Newman et al. (2003) Gupta & Skillcorn (2006) Bond & Lee (2005)
Certainty words	More	Emphasise that what they are sharing is the truth to help conceal their deceptions.	Opinions, feelings Opinions, feelings Psychopaths	Mihalcea & Strapparava (2009) Almela,Valencia-Garcia & Cantos (2012) Almela, Alcaraz-Marmol & Cantos (2015)
Social processes	Fewer	CBCA & RM, lack the real experience to include these nuances.CLH creating these additional details would increase cognitive load.	Psychopaths	Almela, Alcaraz-Marmol & Cantos (2015)

As can be seen from table 1, the linguistic cues associated with deception are contradictory or certain linguistic cues are found to be relevant in some studies but irrelevant in others. Furthermore, they do not always adhere to the direction hypothesised by LBC deception theories. Vrij, Granhag and Porter (2010) argue that the explanation for the inconsistency is simple, there is no 'Pinocchio's nose' of linguistic cues of deception. This is because deception cues are moderated by context and cannot be overgeneralised. Event type, involvement, emotional valence, intensity of interaction and motivation are just some of the moderating variables that have been found (Hauch et al., 2015). Therefore, it's important to create a profile of people exhibiting Munchausen's in order to interpret the relevancy of discriminate cues and their directions in terms of context rather than theory exclusively. This profile includes the moderating variables, motivation, personality, domain specifity and proficiency.

4.2.2.4 MUNCHAUSEN'S PROFILE

4.2.2.4.1 MOTIVATION

People with Munchausen's are likely to be highly motivated to succeed because their aim is to fulfil a need for attention and sympathy. If their deception is uncovered their needs are not met. Woodworth, Hancock and Goorha (2005) found that highly motivated liars using CMC benefit from a motivational enhancement effect, they were more successful at deceiving others. This is because in a CMC environment they have extra advantages. The lack of verbal cues, having unlimited time to create posts, the ability to edit their posts and the being able to engage in selective self-presentation. Research suggests that motivated liars in a CMC environment produce fewer negations and casual terms than unmotivated liars (Hancock et al., 2010). Furthermore, because the motivation to lie for people exhibiting Munchausen's is to elicit attention and sympathy, this motivation may also be reflected in deception cues.

4.2.2.4.2 DOMAIN SPECIFICITY

Linguistic cues of deception also vary according to the topic of the lie. Therefore, linguistic profiles of deception cannot be generalised across topics. Newman et al. (2003) found that their predictive model based on attitude to abortion was more successful within the same topic rather than across different topics. They conclude that there must be a relationship between the topic of communication and the style of

language used to discuss it. This finding is supported by Almela, Valencia-Garcia and Cantos (2013) who reported that different combinations of LIWC dimensions had different prediction success rates according to the topic under discussion. Similarly, Toma and Hancock (2010) found that liars on dating sites included fewer negative words, indicating that they were keener to make a good impression given the context. The topics under discussion by people exhibiting Munchausen's are highly emotive typically revolving around extraordinary scenarios and circumstances full of adversity, in most cases health related. Therefore, the linguistic cues of deception for Munchausen's are likely to reflect the unique style of language associated with it.

4.2.2.4.3 PROFICIENCY

The proficiency of the liar also has a strong bearing on linguistic cues of deception (Vrij, Granhag & Porter, 2010). People with Munchausen's are resourceful and tenacious in their efforts to deceive, due to their deep need for attention and sympathy. Experts believe that they are so adept at lying that most cases go undetected (Feldman, 2004). They have been known to physically induce symptoms, forge medical records, and educate themselves in order to convincingly imitate symptoms, to the extent that they can dupe medical professionals (Pulman & Taylor, 2012). If they are detected and confronted, admissions of guilt are unlikely. Artingstall (1998) argues this is because they believe their lies to be true or they cannot cope with the fallout of admitting to their lies. People with Munchausen's also manage to fool their friends and family so convincingly that even when the ruse is exposed they may not believe it and continue to offer support. In the case of Munchausen by Proxy, if the perpetrator is taken to court, friends and family have been known to offer support by being a character witness for the defence (Lasher & Sheridan, 2004). Given that people exhibiting Munchausen's are so proficient at lying face-to-face, the online environment should be even easier.

People with Munchausen's have a unique profile of features, they are highly motivated by their need for attention and sympathy, they are likely to be high in fantasy proneness, the topics they lie about are emotive and they are proficient liars. Their unique profile means that they are likely to have a unique linguistic profile of deception. Therefore, rather than applying and testing a priori models of linguistic deception all possible linguistic cues and word presence were included in analysis.

Post hoc analysis was then conducted to assess the relevancy of discriminate linguistic cues and their directions in light of the Munchausen's profile presented and existing LBC of deception theories.

4.3 METHOD

Conway and White (2012) describe machine learning, at its very basic level, as teaching a computer about a topic so that it can use this information to perform a task related to the topic, most typically classification prediction. In this way machine learning is distinct from statistics, because statistics teaches humans about a topic so they can use this information to inform their decision making. The "supervised" in supervised machine learning refers to pre-labelling the data with the class it belongs to, as opposed to unsupervised machine learning where the machine must learn to label the data for itself. The learning component of machine learning is the implementation of algorithms that organise the data into their basic constituents and then identify patterns among these constituents. In this way machines can be trained to understand how data is produced and under what circumstances. Once a machine has observed the data and learned from it, a prediction rule can be created and automatically applied to new data to produce a classification prediction. The creation of a classifier involves five stages including, data collection, pre-processing, feature extraction, classifier training and classifier testing. These processes can be implemented by using open source software, the most well-known for text classification is WEKA which was developed by the machine learning group at the University of Waikato (Frank, Hall &Witten, 2016).

4.3.1 CREATING THE DATA SET

The data set was compiled of two types of text. Text written by those with Munchausen's falsely occupying the sick, victim and/or hero role and text written by those genuinely occupying the sick, victim and/or hero role. Cases of Munchausen's were primarily identified by their exposure in communities of online sleuths. These sleuths tried to confirm if there was any evidence in the real world (local news reports, obituaries, and hospital admissions) to support the claims and identity being portrayed by the suspected Munchausen's writer and if the pictures being used were stolen from other users or doctored. More high-profile cases of Munchausen's were identified

through media reports. Although plenty of Munchausen's cases have been exposed, the majority delete their profiles and accompanying entries once the ruse is up. Using the Internet Archive and screen grabs saved by communities of online sleuths, some of text they wrote was collected, however the majority was irretrievable.

It was important to reduce any extraneous discriminant linguistic features by collecting cases of those genuinely occupying the sick, victim and/or hero role that was closely matched with the cases of Munchausen's collected. This reduced the likelihood of the classifier using discriminating linguistic features that were unrelated to Munchausen's. For example, if the cases of Munchausen's collected were all female and the genuine cases were all male, the discriminating linguistic feature would be sex related rather than being discriminating features of Munchausen's. Therefore, the classifier would be extraneous and redundant. To reduce this risk the genuine cases were matched with the Munchausen's cases according to the experiences being discussed and gender of the writers. This is because the linguistic features associated with the experience being written about and the gender of the writer could drown out those of Munchausens if the data sets were biased in either direction. The gender of the writer was categorised into three groups, male, female and male/female. The male/female category refers to a case where both male and female writers were present e.g. a mother and father posting about their sick child under a shared account. In the Munchausen's data set 20% were male, 67.5% were female and 12.5% were male/female. The genuine data set was similar with 18% male, 71.5% female and 10.5% male/female. The experiences written about were grouped into death of child (e.g. miscarriages, cancer), cancer (treatable and terminal cancer), mental illness, abuse survivor (e.g. physical sexual), tragic life changing event (e.g. car crash, murder, disappearance, accident), chronic issues (e.g. chronic pain, cystic fibrosis, paralysis) and heroic roles (e.g. paramedic, soldier). In the Munchausen's data set 24% of experiences related to the death of a child, 19% cancer, 11% mental illness, 8% surviving abuse, 16% tragic life event, 19% chromic issues, and 6% heroic role. Similarly, for the genuine data set 20% of experiences related to the death of a child, 23% cancer, 10% mental illness, 6% surviving abuse, 20% tragic life event, 15% chromic issues, and 5% heroic role. Therefore, neither data set was biased towards a gender or experience which could lead to a classifier using extraneous linguist features associated with either. To achieve this equity, internet searches were conducted to

identify blogs that were closely matched with the gender and experiences being falsely portrayed by the Munchausen's writers. This involved simply running the target experience through the google search engine e.g. 'chronic illness blog' and then collecting amounts of these experiences that were in proportion with those in the Munchausen's data set. The gender of the writer was also factored in to ensure the correct proportions were collected. Authenticity was established if the writer had a transparent public profile (abundance of personal pictures, obituaries, featured in local news, visible occupation, and social circle) compared to those with Munchausen's where there was no real-world evidence to support their claims or identity. The text from both genuine and Munchausen writers was collected from a variety of sources including twitter, Facebook, blogs, autobiographies, and forums. Where possible, extraction software was used including Web Content Extractor, Import.io and the Optical Character Recognition tool for extracting text from screen grabs. However, due to format issues manual collection was also required in some instances. The data set was compiled of 80 cases of Munchausen's who had written text amounting to 1, 573, 957 words and 100 cases of genuine writers who had written text amounting to 3, 066, 634 words.

4.3.2 Stage 2: Automatically extracting the relevant discourse features

4.3.2.1 Pre-processing the data set

Before feature extraction occurs, the data must undergo pre-processing. Sergienko, Shan and Schmitt (2016) refer to this stage as separating the 'signal' from the 'noise' and it involves several tasks. The punctuation was removed from the documents and capital letters were changed to lower case. Stop-word filtering and stemming was also applied using WEKA to try and reduce the number of extraneous cues. Stop-word filtering is the removal of common terms that have little discriminatory power. This helped to free up storage and speed up processing. Determining what words would be removed was based on collection frequency, that is the total number of times a term appeared in the text corpus. Those terms with a high collection frequency were removed as they occur frequently throughout the text corpus and therefore offer little discriminatory value. Stemming was also conducted, and served the same function as stop-word filtering, by reducing inflected words back down to their base form e.g.

banks, banking and banker would be transformed to its base form bank.

4.3.2.2 FEATURE EXTRACTION

Feature extraction involves the texts being represented as feature vectors which allows machine learning algorithms to be applied, because unstructured data is not suitable for computing (Agarwal & Mittal, 2014). Jarvis (2012) describes feature vectors as text-internal features that are countable such as individual words, punctuation, characters etc. The text is transformed into feature vectors which is an ordered set of numbers where every vector number corresponds to the weighted value of a feature, typically relative frequencies of occurrence. Choosing potentially discriminatory feature vectors is arguably the most important step as they are directly responsible for how accurate the predictive model will be and its computational efficiency (Kotsiantis, 2011). Below are the four-primary feature vector extraction methods used in text classification to develop attribute models (Akaichi, 2015; Rice & Zorn, 2003; Boyd & Pennebaker, 2015):

(1) Bag-Of-Words (BOW) – In the BOW's model each word is regarded as a feature and given a weight. The weight can be calculated in several ways the most basic being binary, where a feature weight of one is given if the feature is present and zero if it is absent. More complicated weightings are:

- tf(w) – frequency or number of times a word appears in a document
- df(w) – document frequency or number of documents that contain the word
- tfidf(w) – relative importance of the word within the document.

(2) N-Gram – The BOW's model is the simplest form of feature extraction. However, it excludes several aspects of the text such as word order, syntax and word expressions, these features are picked up by the n-gram model. The n-gram model is a succession of items, typically characters and words, with a sequence length of n. The main types are; unigrams (I / read / a / book / about / cognitive / psychology); bigrams (I read / read a / a book / book about / about cognitive / cognitive psychology); and trigrams (I read a / read a book / a book about / book about cognitive / about cognitive psychology).

(3) Part-of-Speech tagging (POS) – POS involves assigning words a tag that categorises them according to their definition and context. For example, words can be tagged into their grammatical categories (noun, verb, article, adjective, preposition, pronoun, adverb, conjunction, and interjection).

(4) Dictionaries – While BOW's, N-grams and POS models analyse the content and context of texts, dictionaries can be used to analyse the sentiment. Sentiment analysis investigates the opinions, sentiment and emotions conveyed in the text. While sentiment analysis can be conducted by labelling texts according to their sentiment and then developing a classifier to recognise the relevant attributes associated with the target sentiment, there may not be enough training data available. Therefore, a dictionary-based approach is often used instead that contains pre-defined dictionary of words and their corresponding categories. Linguistic Inquiry Word Count (LIWC) (Pennebaker et al., 2007) is a popularly used dictionary analysing text along 80 dimensions, grouped into general descriptors (e.g., words per sentence, percent of word > 6), standard linguistic dimensions (e.g., pronouns, axillary words), psychological constructs (e.g., effect, cognition), personal concerns (e.g., work, health), paralinguistic dimensions (e.g., assents, fillers), and punctuation (e.g., periods, commas).

For the purposes of this study two attribute models were created, LIWC2015 and N-gram. For the LIWC2015 model, the data was run through the LIWC2015 software and the resulting csv output file was converted into an ARFF file that was compatible with WEKA. The N-gram model was created through WEKA by using the StringtoWordVector tool that automatically splits the text up into 1000 N-grams.

4.3.3 Stage 3: Developing statistical classifiers

4.3.3.1 Classifier Training

The fourth stage is applying machine learning algorithms to the attribute models which can learn the discerning feature patterns from the data and use this information to classify new data. Generally, a variety of machine learning algorithms are used and the one that produces the classifier with the best performance is chosen. The most

popularly used machine learning algorithms in binary text classification problems include Naive Bayes (NB), Decision Trees (DT) and Support Vector Machine (SVM) (Bird et al., 2009).

Raschka (2014) describes the NB classifier as a probabilistic learning algorithm that assumes that the presence or absence of a feature within a class is independent of all other features by ignoring the correlations between the features, hence the term Naive. The NB classifier uses the following equation based on Bayes' theorem to calculate the posterior probability.

$$\text{Posterior probability} = \frac{\text{conditional probability.prior probability}}{\text{evidence}}$$

Using spam detection as an example, Raschka (2014) defines the posterior probability as the likelihood that a message is spam based on its feature values. Conditional probability refers to the likelihood of encountering a feature e.g. the word 'offer' in spam. Prior probability is the likelihood of spam by looking at the percentage of message that are spam in the training data and if the training data is a representative sample of the general population. Evidence is the likelihood of a certain pattern occurring independently of the class label. Putting these three elements together the NB classifier can predict the likelihood of a new message being spam by calculating the posterior probability.

Lantz (2015) describes the workings of the DT classifier, which uses a more complex algorithm than the NB classifier. The DT classifier uses a tree structure (like a flowchart) that contains a succession of decisions about features so new data can be classified using predefined rules. The decision tree contains decision nodes (features which require a decisions), branches (indicate the choice made) and leaf nodes (the outcome of a combination of decisions). The decision tree is built using recursive partitioning where the feature values are used to split the data into smaller and smaller strata. The algorithm identifies the feature that is predictive of the class under examination and then divides the feature into values to create the first set of branches. This is referred to as a divide-and-conquer method. This process continues selecting

the next best feature until the tree has outgrown its predefined limit, or there are no more distinguishing features, or most examples at the decision nodes have the same class. A new document can be filtered through the decision tree and the feature values contained within the document are tested according to the decision rules to produce a predicted class.

The SVM algorithm, as described by Sarkar (2016), uses a visual approach by representing the features as points in space and linearly dividing them into their respective classes using the best fitting hyperplane. The class that a new document belongs to is predicted according to which side of the hyperplane its features align. The likelihood of correctly predicting the class of a new document increases as the margins either side of the hyperplane get wider. However, it is rarely possible to separate data linearly and in this case SVM can apply kernel functions that map the features from a 2d view to a 3d view. So now instead of searching for a line that separates the classes a plane is used.

The NB, DT and SVM machine learning algorithms were applied to the two attribute models using WEKA. The classifiers were tested and evaluated using a 10-fold cross validation. 10-fold cross validation was used as Bird et al. (2009) recommends, that if the data set is small, cross validation will increase the accuracy. Cross-validation involves dividing the data into subsets called folds and training the classifier on each fold and then testing it on the remaining folds. By combining the scores from the multiple fold tests, a high accuracy can be obtained. To improve the accuracy of the classifiers even further, attribute selection was also performed in WEKA by running the functions AttributeSelectedClassifier in meta with cfsSubsetEval. These functions are pre-written command lines that are selected and run together in WEKA to reduce the dimensionality of the data before it is moved forward to the classifier. In combination, these two functions choose the best subset of attributes according to attributes that are highly correlated with the class attribute but not with one another. Any redundant and irrelevant attributes were automatically removed to improve the accuracy of the models and reduce the number of attributes required.

4.4 RESULTS

Once the classifier had been trained, the next step involved evaluation of the classifiers performance. The five most commonly used performance metrics are accuracy, precision, recall, specificity and ROC analysis (Tonkin, 2016):

(1) Accuracy is the most basic measure and refers to the percentage of documents that are correctly classified. However, it is important to also examine accuracy in terms of misclassification. That is the frequency of false positives (type I error) and the frequency of false negatives (type II error).

(2) Precision refers to how many of the documents were correctly identified, in other words asking if the classifier is usually correct. Precision is calculated by dividing the number of true positives by the sum of the true positives and false positives.

(3) Recall (also referred to as sensitivity) is the proportion of documents correctly identified in comparison to the overall number of documents that should have been identified. Recall is calculated by dividing the number of true positive by the sum of the true positives and false negatives.

(4) Specificity is the classifiers propensity to avoid making false positive classifications and hence raising false alarms. Specificity is calculated by dividing the number of true negatives by the sum of the true negatives and false positives.

(6) ROC analysis is a graphical plot of the classifiers performance and is used for comparing classifiers. The true positive rate is represented on the Y axis and the false positive on the X axis. The ideal is a true positive rate of 1 and a false positive rate of 0. Therefore, the larger the area under the curve (AUC) is, the better performing the classifier.

4.4.2 RESULTS FROM LIWC2015 MODEL

The dimensions from the LIWC2015 were used as attributes to train the three machine learning algorithms. Table 4.4 displays the results of the three types of classifiers selected, NB, DT and SMO all using 10-fold cross validation.

Table 4.4 Performance of classifiers using LIWC2015 dimensions

Classifier	Accuracy	ROC Area	Confusion Matrix	Classified as: Genuine	Munchausen
NB	81.11%	0.9	Genuine	88	12
			Munchausen	22	58
DT	77.22%	0.79	Genuine	81	19
			Munchausen	22	58
SMV	81.66%	0.81	Genuine	85	15
			Munchausen	18	62

To improve the accuracy of the classifiers, an AttributeSelectedClassifier in meta using cfsSubsetEval was implemented using WEKA. The results and the best subset of attributes obtained are displayed on the following page in table 4.5.

Table 4.5 Performance of best subset of LIWC2015 dimensions

Classifier	Attributes Retained			Accuracy	ROC Area	Confusion Matrix	Classified as: Genuine	Munchausen
NB	Second Person Article Preposition Verb Period	Work Informal Swear Assent Non Fluencies All Punctuation	Social Hear Health Achieve Time Apostrophe	81.11%	0.90	Genuine Munchausen	89 23	11 57
DT	Words per Sentence Second Person Article All Punctuation	Achieve Time Work Informal Swear Assent Apostrophe	Verb Social Male Hear Health	77.77%	0.79	Genuine Munchausen	82 22	18 58
SMV	Words per Sentence Second Person Assent All Apostrophe	Achieve Time Work Informal Swear Punctuation Health	Article Verb Social Male Hear	79.44	0.78	Genuine Munchausen	89 26	11 54

The classifiers with the highest accuracy are NB, SVM and the AttributeSelectedClassifier with NB. The AttributeSelectedClassifier with NB appears to be the best as it could maintain the same accuracy as the NB classifier that used 30 attributes, with only 17 attributes. However, it did not perform as well as the SVM with 30 attributes, but only marginally. To test if this difference between the SVM and AttributeSelectedClassifier with NB were significant, WEKA's Experimenter was used to perform a t-test. The output is displayed below in table 4.6.

Table 4.6 T-test between AttributeSelectedClassifier with Naïve Bayes and SMO

Classifier	AttributeSelectedClassifier with NB	SVM
Accuracy	81.11	81.66

V= Significantly Worse
* = Significantly Better (v/ /*) | (0/1/0)

Using a significance level of $p<.05$, the two classifiers were not shown be significantly worse or better than each other. Therefore, given the smaller number of attributes required by the AttributeSelectedClassifier with NB, it is deemed to be the best classifier obtained using LIWC2015 dimensions. However, it is important to note that it incorrectly classified Munchausen's as Genuine in 27.5% of cases compared to Genuine being incorrectly classified as Munchausen in 11% of cases. Therefore, it is not as accurate at identifying Munchausen's correctly.

4.4.3 EVALUATION OF N-GRAM MODEL

The N-gram model obtained from vectorisation using WEKA was also assessed across the three machine learning algorithms, NB, DT and SVM. The results are displayed below in table 4.7.

Table 4.7 Performance of classifiers using N-gram model

Classifier	Accuracy	ROC Area	Confusion Matrix	Classified as: Genuine	Munchausen
NB	80%	0.79	Genuine	94	6
			Munchausen	30	50
DT	73.33%	0.76	Genuine	77	23

			Munchausen	25	55
SVM	84.44%	0.84	Genuine	91	9
			Munchausen	19	61

As with the classifiers obtained using the LIWC2015 model, AttributeSelectedClassifier in meta using cfsSubsetEval was implemented through WEKA to try and improve accuracy and remove any redundant attributes. The results are displayed on the following page in table 4.8.

Table 4.8 Performance of best subset of n-grams from the N-gram model

Classifier	Attributes Retained					Accuracy	ROC Area	Confusion Matrix	Classified as:	
									Genuine	Munchausen
NB	Began Class Clinic Control Dealing Sense Step	Normal Often Others Painful Place Result Plan	Diagnosis Emotions Energy Future Grateful “” Symptoms	Deep Process Real Recently Learning Memory Run	Grief Imagine Journey	81.67%	0.83	Genuine Munchausen	93 26	7 54
DT	Normal Future	Others Sense	Diagnosis Grief	Plan Clinic		74.44%	0.76	Genuine Munchausen	79 25	21 55
SVM	Began Class Clinic Control Dealing Deep Start	Normal Often Others Painful Place Plan Run	Diagnosis Emotions Energy Future Grateful Grief Result	Step Symptoms “” Real Recently Memory Process	Imagine Journey Learning	80.55%	0.80	Genuine Munchausen	88 23	12 57

Unfortunately, improvement in prediction accuracy only occurred for the NB model. The most promising models were the SVM and AttributeSelectedClassifier with NB, the latter is more desirable as it used a reduced subset of attributes making it more efficient. To test if this difference between the classifiers were significant, WEKA's Experimenter was used to perform a t-test. The results are displayed below in table 4.9.

Table 4.9 T-test between AttributeSelectedClassifier with Naïve Bayes and SMO

Classifier	AttributeSelectedClassifier with NB	SVM
Accuracy	81.67	84.44

V= Significantly Worse
* = Significantly Better (v/ /*) | (0/1/0)

The results show that the classifiers were not significantly worse or better than each other at a significance level of .05. Given that the AttributeSelectedClassifier with NB could achieve a similar result with a reduced subset of attributes, it was deemed as preferable to the SMO which required 1000 attributes. However, like the AttributeSelectedClassifier with NB using the significant LIWC2015 dimensions, it performed relatively poorly at identifying Munchausen's. The confusion matrix in table 6 shows that Munchausen was incorrectly classified as Genuine in 33% of cases and Genuine was incorrectly classified as Munchausen's in 7% of cases.

4.5 DISCUSSION

The two best classifiers obtained using the significant LIWC2015 and N-grams models were both obtained using the AttributeSelectedClassifier with NB. They both had a similar prediction accuracy of 81.11% and 81.67%. However, the classifier obtained using the LIWC2015 model was marginally better at identifying Munchausen's correctly with fewer false negatives. The following sections in the discussion will examine the relevancy of the predictive attributes obtained considering the Munchausen's profile and theories of LBC of deception.

4.5.1 PREDICTIVE LIWC DIMENSIONS

4.5.1.1 CAPTIVATING READERS FOR THEIR ATTENTION AND SYMPATHY

Munchausen users were higher on the dimension, *second person pronouns* (you), which is contrary to the findings of most research studies into LBC of deception (Dzindolet & Pierce, 2015; Hauch et al., 2012; Newman, 2003; Bond & Lee, 2005). The finding does support IDT, that liars use more other-directed pronouns to dissociate themselves from their lies. Taking a closer look however, the lack of consistency i.e. there was no difference on *first person* and *third person pronoun* dimensions, dissociation may not be the only explanation. In the context of Munchausen's profile, specifically motivation and proficiency, *second person pronouns* could be used to appeal to an audience, to get their attention and create an intimate atmosphere (Cui & Zhao, 2014). *Second person pronouns* therefore, may help to fulfil the need which ultimately motivates people with Munchausen's, to elicit attention and sympathy through their deception. *Second person pronouns* would draw people in to read their stories and make them feel emotionally involved enough to offer support to the person with Munchausen's.

In the same vein, the dimension *informal speech* was predictive of Munchausen's, specifically more *swear words*, *assent*, and *non-fluencies*. No consistent association between the dimension *informal speech* and deception has been made by previous research. Burns and Moffitt (2014) did find assent was associated with deception. But in the specific context of emergency calls, it was used to limit and control answers ('yes'), making its generalisability to Munchausen's spurious. However, Zhou et al. (2004) offer another explanation, by hypothesising that informal language would help build rapport and familiarity, making it appear that the liar was not manipulative. They found that increased informality, typographical errors measured this, was associated with liars. Again, in the context of Munchausen's motivation, *informal speech* would help in befriending others, with the purpose of exploiting this friendship for attention and sympathy. This strategy also demonstrates the proficiency of people exhibiting Munchausen's.

Once their need for attention has been gained through cultivating immediacy, they need to maintain it by continually providing a slew of drama, most typically in the

form of "pseudologia fanstastica". This may explain why people exhibiting Munchausen's were also not as high as genuine writers on the dimension *achieve*. Their modus operandi is not to leave the sick role and move forward, instead constantly oscillating between death bed and miraculous recovery. This is a key feature in recognising MbI, as well as the creation of dramatic events that intensify as attention is shifted away from them (Feldman, 2000).

4.5.1.2 SOCIAL OUTLET

Although health issues are the primary means through which people exhibiting Munchausen's try to elicit attention and sympathy, surprisingly, genuine writers were higher on the *health/illness* dimension of the LIWC2015. However, people exhibiting Munchausen's were higher on *social processes* which denotes human interactions. One possible explanation for this discrepancy may highlight a key linguistic feature of Munchausen's related to domain specificity. They may lack the vocabulary to discuss *health/illness* because they never experienced it, which would support RM and IDT, and/or they may be more interested in the *social processes* that accompany it rather than the medical aspect. For example, Lawlor and Kirakowski (2014) found that people with Munchausen's experienced social isolation which prompted them to seek affection by falsely occupying the sick role. Therefore, MbI goes beyond just falsely occupying the sick role it involves the need for interaction. A common feature of MbI is 'sock puppets', creating people who don't exist. Their inclusion could be a way of alleviating social isolation and explain the higher frequency of social processes words. In one case of MbI, 71 sock puppet accounts on Facebook were created. While it's believed that the primary purpose of sock puppets is to add credence to the story, it could also be a roleplaying exercise to help alleviate social isolation. For example, adults who have experienced childhood abuse (which is common among people with Munchausen's) have been known to create imaginary friends and/or family enabling them to have quasi-relationships (Sanderson, 2006). The increased use of *social processes* words may not be exclusively related to 'sock puppets' which have their own profiles and voices. It could also be the person exhibiting Munchausen's creating a fictitious circle of friends and family within their own fictional narrative so they feel less isolated and more cared for. While this is one plausible explanation for the dominance of *social processes* it is by no means the only one. An alternative explanation can be found in VI theory and relates to the criteria denotative specificity

(Weiner & Mehrabian, 1968). The increased usage of words pertaining to *social processes* by people exhibiting Munchausen's may be indicative of over inclusion. A technique associated with deception that indicates low verbal immediacy as the deceiver avoids being overly specific by using broad terminology. By relying on words associated with *social processes*, in particular third person pronouns, the person exhibiting Munchausen's may be using over inclusive terminology to refer to specific people.

4.5.1.3 RELATION TO LBC THEORIES OF DECEPTION

The LIWC2015 dimensions associated with Munchausen's did overlap with LBC's theories of deception, particularly those relating to cognitive complexity. Consistent with IDT, CLH and RM, people exhibiting Munchausen's appeared to lower cognitive complexity. Like previous research that used LIWC dimensions, verbs were predictive of people exhibiting Munchausen's by providing simple descriptions avoiding more complex evaluations and judgements (Bond & Lee, 2005; Hauch et al., 2012; Newman et al., 2003; Gupta & Skillcorn, 2006). Conversely, fewer prepositions were also predictive of people exhibiting Munchausen's. Prepositions belong to a class of words called exclusive words, denoting what did or did not happen, thus increasing the cognitive load (Toma & Hancock, 2012). The decreased use of exclusive words has been found to be associated with deception (Bond & Lee, 2005; Hauch et al., 2012; Newman, 2003). Similarly, the use of articles was comparatively lower than genuine writers. Articles are associated with linguistic complexity and trustworthiness because they are linguistic markers of concreteness, well known and easy to comprehend (Toma & Hancock, 2012). In contrast, the LIWC2015 dimension *punctuation*, specifically *periods* and *apostrophes*, was found to be high among people with Munchausen's. This conflicts with IDT and CLH, as punctuation increases sentence complexity and therefore the cognitive load (Zhou, 2004).

Genuine users were higher on the dimensions *time* and *work*. Liars using fewer perceptual words supports the meta-analysis conducted by Hauch et al. (2012). It also supports CBCA, RM and CLH which posit that liars lack the externally derived memories required to include relativity words and that the attempted inclusion of descriptive details would increase cognitive complexity. However, this contradicts with the findings that people exhibiting Munchausen's were higher on the dimension

hear. Although Hancock et al. (2008) did find that deceptive online daters did use more sense words, arguing that they might be trying to increase credibility. This could also apply to people exhibiting Munchausen's who are known to be proficient liars. With regards to the inclusion of more detailed information increasing cognitive complexity, genuine writers were also higher on the dimension *work*. This may reflect the practical aspect of being ill, something people exhibiting Munchausen's would not experience.

4.5.2 PREDICTIVE WORDS IN THE N-GRAM MODEL

Considering there are no theories or previous research which elucidate the use of specific words and their relevancy to deception, it makes it more difficult to discuss the context of the predictive words obtained from the N-gram model. All the predictive words were more frequently found in the text written by genuine writers. Therefore, to gain a better understanding of their relevancy they were analysed using the Summary Variables category of the LIWC (see table 9). According to Pennebaker et al. (2015) the Summary Variables are research-based composites, converted to a 100-point scale where 100 is very high on the dimension and 0 is very low. With regards to the dimension *Emotional Tone*, above 50 suggests a positive tone and below 50 suggests a negative tone. The Summary Variables category was specifically used as it does not use word frequency which would be inappropriate for the analysis of a list of words. The four dimensions included under Summary Variables are *analytical thinking*, *clout*, *authenticity* and *emotional tone*.

Table 4.10 LIWC dimensions of the subset of predictive words from N-gram model

Summary Variable	Score
Analytical Thinking	83.9
Clout	36.9
Authenticity	99.0
Emotional Tone	25.8

The subset of predictive words, which occurred more frequently in genuine texts, were very high on the *Analytical Thinking* and *Authenticity* dimensions. The predictive

words were associated with less of a narrative style (focuses on the here and now, and personal experiences) and more of an analytical approach (identifying conceptual categories and organising them into hierarchies) to issues being experienced (Jordan & Pennebaker, 2016). As well as the use of a more authentic tone, where language is more humble, vulnerable, and personal (Newman et al., 2003). In contrast, the subset of predictive words which occurred more frequently in genuine texts were low on both Clout and Emotional Tone. Meaning that the predictive words were not highly associated with social status, confidence or leadership and had a negative emotional tone (Cohn, Mehl & Pennbaker, 2004; Kacewicz et al., 2013).

It would be inappropriate to extrapolate the linguistic dimensions of the subset of predictive words as distinguishing features of genuine writers in comparison to those exhibiting Munchausen's. Particularly as none of these dimensions were found to be predictive attributes across the full texts (see table 3). But it does provide context for the subset of predictive words obtained from the N-gram model.

4.6 LIMITATIONS

An important limitation of the study that should be noted relates to the generalisability of the data collected. The text collected from those exhibiting Munchausen's and genuine writers may not be wholly representative of both groups. This is because the collection of reliable data necessitated that those exhibiting Munchausen's had to have been exposed and evidence of an offline presence was required to confirm genuine writers. Therefore, more overt and 'successful' cases of Munchausen's and genuine writers who wish to remain anonymous are not represented in the data collected. Unfortunately, this limitation could not be negated because some form of confirmation was required to ensure the quality of the data for classification. A second limitation is the unwitting inclusion of biased data. That is either the Munchausen's or genuine data sets not having proportionate amounts of data from a specific cohort resulting in the discriminate linguistic cues being extraneous. To avoid this, proportionate amounts of data were collected in both data sets according to gender and the experience being written about. However, other cohorts may have been unwittingly included resulting in lack of proportional data. These include cohorts that were difficult to discern, particularly for the Munchausen's data set, including cohorts categorised by age,

location, educational attainment, occupation etc. These cohorts may use distinguishing linguistic features, and if there were unproportionate amounts included in either the Munchausen or genuine data set, the classifier would simply be using discriminate features associated with the dominant cohort rather than Munchausen's. While this is a potential limitation that needs to be borne in mind when interpreting the relevancy of the discriminate linguistic cues, it cannot be totally ameliorated and is an inherent risk in text classification research. However, every effort was made in ensuring that the cohorts that could be discerned did not dominate either data set. Furthermore, the methods used to authenticate genuine and Munchausen's cases were robust so that the target cohort dominated both data sets. This means that even if a cohort was unwittingly unproportionate in either data set, the target cohort would still dominate the data so that the most consistently present, and therefore discriminate linguistic features, would be related to the target cohort i.e. Munchausen's. Furthermore, the discriminate linguistic cues associated with Munchausen's did fit theoretically and this may be further evidence that additionally cohorts did not unwittingly dominate the data. Despite the low risk of the discriminate cues being associated with a cohort other than Munchausen's, it is still important to acknowledge this potential limitation given the implications it would have for the validity of the classifiers.

4.7 CONCLUSION

The SLP paradigm was successfully applied to develop two viable classifiers. The classifiers were particularly adept at detecting text written by genuine writers, using a N-gram model and a model based on the LIWC2015 dimensions. It is hoped that they could be of practical value to moderators in identifying possible cases of MbI. But ultimately, given the high false negative rate in identifying text written by people exhibiting Munchausen's, online sleuths are still needed as up to a third could slip through the net undetected. More research is required to reduce the rate of false negatives. Collecting a larger corpus and looking at the interactions between people exhibiting Munchausen's and the people they are trying to dupe, because as IDT notes the relationship is dyadic.

The classifier derived from the LIWC2015 dimensions is more theoretically valuable because it contributes to pre-existing psychological theories of deception. Whereas

extrapolations cannot be made from the subset of predictive words obtained from the N-gram model, a framework for understanding their linguistic makeup can only be applied. The discriminate LIWC2015 dimensions obtained support the growing view within the theories of deception that there is no one-size-fits-all model for LBC's of deception. Domain specificity must be considered when interpreting cues of deception, as some may not be consistent with the overarching theoretical LBC's of deception but are specific to the given context. In support of this view, the study uncovered some interesting strategies employed by people exhibiting Munchausen's which betray their deceptive behaviour, but which cannot be understood in terms of theories of LBC's of deception alone, but in the context of the profile of Munchausen's. The profile shows that motivation to receive attention, and maintain this attention, as well as the proficiency at lying could be associated with increased use of second person pronouns and informal language and comparatively less use of achievement related words. With regards to domain specificity, the expected focus for people exhibiting Munchausen's was health/illness but in fact it was more frequently discussed by genuine writers. Instead words relating to social processes were more frequently used by people exhibiting Munchausen's. One possible explanation explored is that this may be indicative of their need to alleviate social isolation, a possible underlying motivation for their behaviour. The remaining discriminant cues obtained did support LBC's of deception theories, particularly those associated with reducing cognitive complexity which specifically supports IDT, CLH and RM. While pre-existing theories of LBC's of deception are a valuable supplementary tool in interpreting the relevancy of cues of deception, this study supports the growing view that cues of deception also need to be interpreted in terms of their domain specificity.

CHAPTER 5.0 CONCLUSION

5.1 INTRODUCTION

The suite of studies presented in this thesis were conducted to address a limitation of Lawlor and Kirakowski's (2014) study, with the aim of acquiring knowledge about FD and MbI that could be used to inform the development of a method to detect MbI and to learn more about the dynamics of online communities in the face of MbI. A method of detecting MbI is required to help users from being emotionally and financially exploited as well as an attempt in safeguarding the space facilitated by the unique features of online communities that foster personal empowerment. In addition, a method of detecting MbI is a useful vetting and investigative tool for internet mediated researchers. Before a method of detecting MbI could be chosen and developed, two exploratory groundwork studies were required. This was because of the poor understanding of FD and MbI, primarily due to the lack of studies that use first-hand accounts. Previous research has been monopolized by observational case studies which threw up numerous theories that were largely speculative and had little practical application value. The research undertaken for this thesis expanded existing knowledge and produced a practical application value to the research. This was achieved by using first-hand accounts of those directly affected by FD and MbI, FD sufferers and online users who have encountered MbI.

The decision to develop a text classifier through machine learning to detect MbI was based on the information garnered from *paper 1; Munchausen by Internet, paper 2; When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder, and paper 3; Claiming someone else's pain: A Grounded theory analysis of online community user's experiences of Munchausen by Internet.* Although FD sufferers were found to be aware and perturbed by their behaviour, the addiction element to it meant that they would likely try to evade direct testing and would be reluctant to participate in the development of such a test that would ultimately be used to stem their addictive behaviour. Using cues as a checklist to identify MbI was also problematic as such cues are ambiguous and could be open to biases in human judgement. Therefore, a more covert method that could circumnavigate these issues was required, a text classifier. The collection of the data required to train the classifier and the implementation of testing did not require participation from MbI users. The

discriminate attributes used by the classifier are unambiguous and decisions of their absence/presence are made by a machine which is not subject to human biases. The most suitable method for detecting MbI was therefore deemed to be a text classifier which was developed in *paper 4; Detecting Munchausen by Internet: Development of a Text Classifier through Machine Learning.*

Although the findings from the studies were used to develop a method to detect MbI, which has a practical application value for researchers and moderators, their reach went beyond this. They also made theoretical contributions to several disciplines and paradigms within psychology, namely Psychiatry (new theories of FD), Cyberpsychology (dynamics of online communities in the face of MbI), Social Psychology (LBC's cues of deception) and the SLP paradigm (successful application to MbI detection). The remainder of the conclusion will enlarge on the contributions made by the research conducted for this thesis, both theoretical and practical, discuss the limitations of the studies and provide directions and areas for future research.

5.2 PRACTICAL APPLICATION VALUE

The development of a viable classifier to detect MbI members from genuine members has a practical application value for moderators of online communities and internet meditated researchers. The classifier could be used by moderators of online communities to vet the posts written by its members. However, while the classifier may identify approximately 70% of suspected cases of MbI correctly, given the rate of false negatives at 30% and the cost of falsely accusing an already vulnerable person of lying, it would be imprudent to use the classifier for confirmation of MbI. The classifier should only be used in place of high-level clues of deception which were previously used to detect possible cases of MbI and were at greater risk of being applied to genuine members of online communities. Ultimately, online sleuthing is still required to collect authoritative clues to confirm the suspected case of MbI, it is the only fool-proof method of positively identifying MbI. It is hoped that the classifier would contribute to safeguarding online communities, not only from being emotionally exploited by MbI users, but also by helping to alleviate the air of suspicion in online communities. The air of suspicion that has the potential to erode the ability of online communities to foster personal empowerment, by offering a risk-free space

that is non-judgmental to confer normality and cultivate interpersonal support, particularly around identities, illnesses and experiences which are marginalised or stigmatised. The classifier could also help to combat the financial exploitation perpetrated by malingerers in online communities that provide fundraising platforms. By detecting possible cases of fraud that require further investigation before the malingerer has a chance to disappear.

The original impetus for a method of detecting MbI that was derived from the study by Lawlor and Kirakowski (2014). A limitation of their research was the possibility of unwittingly including participants who had MbI or OSG's that were overrun by MbI users in their research. This is a limitation which internet mediated researchers need to be more cognizant of as it has the potential to make the findings of research spurious. Particularly for internet mediated researchers who are recruiting participants from or analysing OSG's where MbI is likely to be more prevalent. The classifier has the potential to be the de facto standard for filtering potential participants or online communities so that cases of MbI are removed during data collection to safeguard findings. The classifier could also be used as an investigative tool in its own right rather than just for vetting participants. For example, to explore the prevalence of MbI in online communities.

5.2.1 LIMITATIONS AND FUTURE DIRECTIONS

While the viable classifiers developed had an overall performance accuracy of approximately 82%, the rate of false negatives when identifying MbI is problematic. However, given that the current method of using high-level deception cues is problematic (ambiguous and vague cues, must be identified within the narrative in an unsystematic way and open to human judgement bias) the classifier offers a superior alternative free of these limitations. The classifier is also more cost sensitive because it errs on the side of caution when identifying cases of MbI. This is because there were more false negatives (30%) when identifying MbI and less false positives when identifying genuine writers (10%). In other words, the classifier is less likely to falsely identify a genuine writer as being MbI and instead more likely to identify MbI as being genuine. Given that it would be costlier to falsely accuse someone who is genuine as lying compared to falsely 'accusing' an MbI user as being genuine, the classifier is

cost sensitive which goes in its favour. However, the accuracy in positively identifying MbI could still be improved. The most obvious way is to collect more data. This is not likely to be easy as exhaustive searches have already been conducted for this study and one of the primary barriers was that once a case of MbI was exposed their postings were deleted, even with the aid of the Internet Archive they were frequently irretrievable. One possible option would be to try and reach out to those who had been deceived to share their personal correspondence. They might not however be very forthcoming given the sensitivity of the topics that might be under discussion. Another option would be to include the dyadic relationship involved by analysing the responses to the MbI users. The classifier may be able to notice subtle differences in correspondences between MbI users/genuine and genuine/genuine that could be used to identify differences in the conversations which are indicative of MbI.

There is also issues around the representativeness of the sample used to develop the classifier. Identifying cases of MbI meant that they had been exposed and therefore represent unsuccessful cases of MbI. There could be covert cases of MbI that are never exposed and the classifier may not be able to notice these cases if they are linguistically compatible with genuine writers. However, given that LBC's of deception are not used in human judgements of MbI they may share similar LBC as the sample used in this study but are better at disguising observable cues. The sample of genuine writers used had to have an offline presence to confirm they were genuine. Therefore, genuine writers who wished to remain anonymous could not be included. This means that there could be an overlap between MbI and anonymous genuine writers which would not have been incorporated into the classifier. Given the need to ensure the quality of the data, genuine was genuine and MbI was MbI, there was no way of overcoming this limitation without compromising the quality of the data, which is of paramount importance.

5.3 THEORETICAL IMPLICATIONS

5.3.1 NEW THEORIES OF FD

Psychiatry's understanding of FD has been constrained by the lack of first-hand accounts of the disorder. This had led to diagnostic criteria that have little or no practical value to clinicians and a lack of successful treatment strategies as FD

sufferers are perceived as being disinterested. Discouraging FD sufferers to seek help in the first place or engage in treatment programs over the long-term. Perhaps more importantly, it has resulted in the disfranchising of people who suffer from FD as oblivious and having no interest in recovery. This disfranchising may in part account for the stigma which surrounds the disorder and the reluctance of sufferers to seek professional help or engage in research that would better our understanding of the disorder and help inform treatments with better outcomes. The findings from *paper 2; When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder* impinge on psychiatry's traditional understanding of MbI with the hopes of stimulating new research which will overlay traditional theories grounded in speculative observational studies with new innovative theories that incorporate the experiences of those suffering from FD.

The findings from this thesis supported the traditional view that the motivation for FD is multifaceted and there is no one single motivating factor for the behaviour, other than to say that the aim is to seek out attention and support to fulfil an emotional deficit. What creates this deficit would appear to be contingent upon the individual's personal circumstances. However, the findings did conflict with the remaining traditional views that motivation was unconscious, there were no internal negative symptoms experienced and hence no impetus for recovery. In contrast, members of the OSG for FD were spontaneously discussing why they behave the way they do, indicating that they are aware of possible motivating factors for their behaviour. They described a long list of negative symptoms they experienced personally (poor mental health, addiction, dissociation from real self, negative self-perception, reduced quality of life) and as a direct consequence of deceiving others (guilt and shame, isolation, ostracised). These negative symptoms provided the impetus for recovery, which traditional theories believed did not occur for those with FD as they were oblivious. Recovery however, did not entail seeking professional help, this was not due to disinterest as was previously assumed, but instead due to a variety of fears around disclosing FD to a therapist, losing family/friends once they found out about the deception, fear of confronting the underlying issues which precipitated the behaviour, logistical difficulties, fear of recovery and fear of perpetuating FD by seeking help. These fears led to a dependence on self-management strategies to try and recover or stem the behaviour by themselves, which further highlights their desire to recover. A

variety of strategies were employed including confronting the precipitating issues for FD, occupying the hero role rather than the sick role, self-imposed isolation, changing their environment, fear of re-incarceration, replacing the negative coping mechanism of FD with a positive one, disclosing FD to prevent them getting away with enacting it, self-insight into the behaviour, fear of being found out, forming new relationships, communicating their needs, having dependents, correcting their lies immediately, having a good support system, applying the 12 step program to FD, maintaining healthy boundaries, needing to disappear at times, implementing positive changes and using logical thinking. The variety of strategies used underscores the variety of motivations for the behaviour.

Beyond challenging the traditional psychiatric theory of FD, the findings also attempted to create a new theory that incorporated these new findings. The theory suggests the inclusion of FD as a subcategory of 'Addictions and related disorders' in the DSM. This would help to overcome the practical limitations of the current diagnostic criteria of FD, the poor treatment strategies available (applying treatments for addiction might be beneficial) and may help in de-stigmatizing FD by viewing it primarily as an addictive behaviour rather than a violation of the sick role. This suggestion is tentatively made because of the common thread which ran throughout the findings, the overlap between these new characteristics of FD and addiction, specifically Orford's (2001) excessive appetitive model. Orford (2001) defined addiction in terms of the emotional rewards associated with a behaviour that serve an emotional function. As applied to MbI, the emotional reward is occupying the sick role for attention/affection and the emotional function is to fill an emotional void. There are parallels between the negative symptoms associated with cognitive dissonance, which is experienced by those with addiction, and the negative symptoms experienced by FD sufferers. Cognitive dissonance occurs because of conflicting emotions. On the one hand, the positive emotions associated with enacting the addictive behaviour and on the other, the negative emotions of engaging in a behaviour that is negative. In the case of FD sufferers, it results in a slew of negative internal symptoms associated with demoralisation. There are also strong parallels between FD and addiction with regards to the recovery process, sharing similar barriers to seeking help and a reliance on self-recovery strategies to try and control the behaviour.

The overlap between FD and addiction had implications for the development of a method to detect MbI. It meant that MbI users were unlikely to be very forthcoming to participate in any testing that would try to uncover traits associated with their personality that could be uncovered through psychometric testing. This is because if such a test was developed it would impinge on their ability to enact FD online without going undetected. This would mean that they could not satisfy their addiction to the attention and sympathy of the sick role as freely as before. Furthermore, given the fears around being exposed as having FD, MbI users may be just as reluctant to participate in research as their offline counterparts. Therefore, a method which could be developed without the direct participation of MbI users was required. This was further elucidated upon in *paper 3; Claiming someone else's pain: A Grounded theory analysis of online community user's experiences of Munchausen by Internet*, where the text classifier was deemed to be best method.

5.3.1.1 LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

The study conducted to produce these new theories of FD is not without limitations. One of the most important limitations that needs to be considered is that the sample may not be representative of FD. The sample in this study, by definition of seeking out an OSG for FD, are self-diagnosed, self-aware and have a desire to recover. They may only represent a sub-category of FD and as such be the exception rather than rule. Therefore, the study may have identified a sub-category of FD to which these theories apply rather than these theories being applicable to all FD sufferers. Despite this limitation the study certainly should give pause for thought. It is the first large scale study of first-hand accounts of FD and therefore has advantages over previous studies. It also offers a new perspective on FD that has the potential to overcome the diagnostic, treatment and disfranchising issues associated with our current understanding of FD. This study has opened the door to reconceptualising FD in a more pragmatic way. However, far more studies are required using a more diverse sample of FD sufferers to support or reject the theories presented. This study has also offered a new way of tapping into FD sufferers who are difficult to find and reluctant to participate in studies. By recruiting them online they are spared any fears around disclosure due to their anonymity and therefore maybe more forthcoming. Unfortunately, such studies will always be limited by questions around whether a sample which willingly engages in research are truly representative of a cohort of

people who were previously believed to be aloof. Despite this, given the stagnation and our lack of pragmatic understanding of FD, new research needs to be conducted to move it forward.

By extension, the findings could also be applicable to MbI. However, it may be that MbI is its own unique sub-category within FD with its own unique aetiology and motivation. There is anecdotal evidence to suggest that some cases of MbI are simply people with FD using the internet as another means of enacting FD for additional attention/affection or as way of adding credence to their FD by continuing it online. This is because if someone is genuinely sick offline they are likely to share their illness online too through social media accounts. For those who enact it only online, the motivation and aetiology could be unique or they may use the internet as a 'safe' means of dealing with their desire to enact FD which would otherwise impinge on their daily lives. Either way, research needs to be conducted to untangle whether MbI is unique or an online extension of FD.

5.3.2. DYNAMICS OF ONLINE COMMUNITIES

Cyberpsychology research has not focused on the dynamics of MbI within an on online community. In fact, MbI has not received much research attention, despite its potential to interfere with the efficacy of online communities and in particular OSG's. Given that the efficacy of OSG's as a therapeutic aid is an important research area because of their growing popularity, not just among peers, but also for mental and physical health organisations, it is surprising that the impact of MbI has been overlooked. One possible explanation is that people are simply not aware of MbI, especially given the incredulity of someone pretending to be terminally ill or having a terminally ill child. The anecdotal evidence indicates that MbI may be more commonplace in OSG's especially as it offers a risk-free environment to perpetrate deception.

The study conducted for *paper 3; Claiming someone else's pain: A Grounded theory analysis of online community user's experiences of Munchausen by Internet* was an exploratory study to investigate the dynamics of MbI within an online community. Only one study by Feldman (2000) had examined the impact of MbI within an online community but this was based on four cases studies and only reported the observable

impact once the ruse had been exposed. The aim of this study was to extend on Feldman's (2000) work by including a larger sample and analysing the first-hand experience of those who have encountered MbI. The findings showed that MbI users were specifically targeting the occupation of 'ideal victim' persona. This persona offered numerous benefits because it elicits the most amount of sympathy, the legitimacy of the person occupying the persona is not questioned given their vulnerable status (making it easier for MbI users to successfully deceive others) and matching the target community with the persona ensures the most amount of sympathy possible is elicited. While it was anecdotally known that MbI users preferred pretending to be expectant mothers with complicated pregnancies or parents of a sick child, applying the 'ideal victim' persona theory explains why they have a preference for occupying these roles. A surprising finding from the study which conflicts with the sentiments of those who have been duped by FD and the stigma surrounding the disorder, was that although online community users felt exploited and manipulated, the majority were sympathetic towards those with FD. Although this would appear to be heartening for those with FD who maybe be struggling with self-stigma, this disparity could be accounted for by the lack of physical proximity that would minimise the factors which contribute to a stigmatised attitude.

The second impetus for conducting the research for this thesis was to explore how MbI impacted on the dynamics of an online community. This was because of the study by Lawlor and Kirakowski (2014) that noted a limitation of their research in terms of the unknown impact of MbI on the efficacy of an OSG's. This also has wider implications for other efficacy studies of online communities which should acknowledge the potential of MbI to interrupt the therapeutic benefits of OSG use. The findings show that the presence or even the perceived presence of MbI results in self-censorship and reserved participation, out of the fear of being duped or being falsely accused of being a fake. This in turn has the potential to limit opportunities to confer normality and cultivate interpersonal support. This is because, the perceived need for self-censorship and reserved participation, belays a lack of interpersonal trust, which is required for initial participation, development of supportive relationship and the creation of a non-judgmental atmosphere where difficult and a-typical emotions, experiences, issues, and identities can be shared, without being accused as fake. This allows the community to confer communal normalisation to a-typical issues and identities characterised by

stigma. Instead some users felt they must replicate the common narrative of the community offering no opportunity for growth and personal empowerment. Self-censorships lead to extraordinary but true personal details being omitted out of the fear of being falsely accused of being fake. This leads to the creation of "conventionalised" narratives of identities and issues which offer no room for deviations and may in fact not be truly reflective. This is perhaps the biggest drawback of MbI for OSG's that was previously unreported because of its subtlety. OSG's have been championed because of the unique risk-free environment that allow for personal empowerment around marginalised and stigmatised identities and issues, a space that isn't always available offline. While MbI has the potential to erode the personal empowerment facilitated by OSG participation, it did not appear to have a completely deleterious impact on online communities. This is testament to the resiliency and importance of online communities to their members. Given the potential impact on personal empowerment of the presence or perceived presence of MbI in online communities, it's important that researchers take this dynamic into account when assessing the efficacy of online communities and moderators (particularly those of online communities belonging to organisations) need to be pro-active in trying to control and reduce the atmosphere of suspicion which is the catalyst for self-censorship and reserved participation.

With regards to detecting MbI, the remainder of the findings focused on the identification and management of MbI. This information was particularly valuable in making the decision as to what option was most appropriate for developing a method of detecting MbI. The findings support the observable cues associated with MbI that were recognised by Feldman (2000), dramatic/ a-typical stories and contradictions/inconsistencies. These are regarded as high-level cues which raise suspicion. However, the problem was that the online community users believed these were vague and ambiguous and therefore could be easily applied to a genuine user, even themselves. The vagueness of cues embedded within the narrative and the difficulty of quantifying them led to the conclusion that LBC's of deception would be a more appropriate method of detecting MbI, which led to the idea of a text classifier that could be automatically implemented. In addition to high-level cues of deception, online community users typically used authoritative reference to confirm their suspicions. While the text classifier could replace the use of vague and ambiguous

cues embedded within the narrative, nothing can replace authoritative references in incontrovertibly identifying MbI. The authoritative references used included identifying duplicate IP addresses, investigating if the person existed offline and checking if identities had been stolen. Due to the laborious process involved in uncovering MbI and the lack of sanctions, many online community users believed that preventative strategies were the most beneficial, primarily maintaining healthy boundaries and not becoming emotionally involved. While this is certainly prudent advice, lack of emotional involvement feeds into the erosion of online communities fostering personal empowerment. The study highlighted the previously undocumented concern that OSG's users have about MbI encroaching into their online communities and the strategies that they employ to detect it and limit its impact.

Overall, the study gives an insight into the dynamics at play in an online community which is coping with the presence or perceived presence of MbI in their online community. The effect of MbI is clearly significant and its effect on the dynamics of an online community should be factored into efficacy studies. Furthermore, the study underscores the importance of a classifier, not only for research purposes, which was the impetus for its development. But also for moderators as a replacement for the current high-level cues that are used, cues embedded in the narrative that are ambiguous and vague.

5.3.2.1 LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

The representativeness of the sample is a possible limitation of the study. While the sample size is relatively big there are some concerns around whether their opinions and experiences are truly representative of online community users. They maybe more exercised about MbI by fact that they have engaged in a discussion about MbI. Therefore, most online users may not be particularly perturbed by MbI and as such would not engage in discussion around it. Another issue with the sample pertains to whether MbI users were also within the sample. This may in part explain why the online community members appeared to be so sympathetic towards MbI and provided theories regarding the motivation for the behaviour which drew parallels with those found in *paper 2; When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder*. Unfortunately, neither of these limitations were avoidable, and considering the difficulty in collecting first-hand accounts of MbI and

the lack of research into the dynamics of MbI within an online community, the research was still warranted.

Given the potential for MbI to erode personal empowerment, more research is required to confirm or reject this. An interesting study to assess this could use the classifier to estimate the prevalence of MbI within an online community and then compare this to measures of personal empowerment among members of the online community. Outcome measures beyond personal empowerment could also be included e.g. measure of therapeutic outcomes. This further highlights the usefulness of the classifier as an investigative tool for internet mediated researchers, not just for vetting participants. More research is also required into methods of preventing MbI through deterrence, this was a concern raised by online community members who believed there was nothing to stop the behaviour. While some suggestions have been made to discourage trolls, the same could be applied to MbI, including making participation more onerous (requiring registration, verification system, probation period) and accountability (publicly displaying IP address, posting bonds). Publicly announcing that a classifier to detect MbI is being used within the community may also act as a deterrent. Other options also need to be explored by new research and more needs to be done to assess if MbI should be regarded as a cybercrime that should be legally pursued. This is particularly relevant for cases involving the theft of photos or even complete online identities and where medical misinformation is being purposely spread endangering the lives of others.

5.3.3 USING THE SLP PARADIGM TO IDENTIFY LBC'S OF DECEPTION

The research conducted in *paper 4; Detecting Munchausen by Internet: Development of a Text Classifier through Machine Learning* not only had a practical application value, it also had theoretical implications. It supports the use of the newly developed SLP paradigm (Hancock, 2010) in identifying online identity deception. The framework provided by the SLP paradigm provides a systematic method for identifying and extracting potential discriminate attributes and then using these to train a classifier. Before this paradigm the use of text classifiers was dominated by computer science research that did not consider the theoretical significance of the LBC's that were identified by the classifier, instead the performance accuracy was the bottom

line. The SLP paradigm allows for an interdisciplinary approach which incorporates both psychology and computer science. By using the SLP paradigm the study could go successfully beyond just developing a classifier using computational techniques, by integrating it with psychology theory to show how the discriminate LBC's of MbI are situated within the wider context deception theory. The use of the LIWC2015 model to develop the classifier was pivotal to making sense of the attributes as the N-gram model was of little theoretical value beyond categorizing the set of discriminate words, whose results could not be extrapolated to theory.

The discriminate LBC's obtained from the LIWC2015 supported pre-existing theories of LBC's of deception. Specifically, those relating to reducing cognitive complexity, this supports IDT, CLH and RM as well as numerous studies which have been conducted into online deception that utilised the LIWC. However, given the importance of domain specificity in interpreting LBC's of deception and the associated inconsistency in universal LBC's of deception, arguably the most valuable and perhaps consistent cues indicative of MbI are those associated with the Munchausen profile and are therefore domain specific. These included cues indicative of the Munchausen personality, their motivation for attention and to maintain this flow of attention towards them as well as their proficiency in having these needs met (more second person pronouns, more informal language and less achievements related words) and the underlying need that this attention quenches, alleviation of social isolation (lower use of health/illnesses related words and a higher use of social processes words). So, while the study supports existing theories of LBC's of deception it also supports the growing belief that LBC's are domain specific and as such there is no LBC of deception that equates to 'Pinocchio's nose'. The study has also created the first linguistic profile of MbI that provides a starting point for future research.

5.3.3.1 LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

The same limitation of the representativeness of the sample used, which was discussed in relation to the development of the classifier, is also relevant here. The LBC's associated with MbI may only be specifically applicable to cases of overt MbI which has been exposed. This limitation is unavoidable because collecting MbI data, by definition, necessitates that the user has been exposed. There is also that possibility that both overt and covert MbI users share the same linguistic markers but covert MbI

users are more adept at concealing observable cues that are visible to human judges. The LBC's of MbI need to be investigated by conducting more studies to support or reject the profile that was found by this study. However, a major obstacle for these studies will be collecting more data, as mentioned, this study conducted an exhaustive search. The only remaining option is to request those who have been in direct contact with MbI users to share any correspondence they may have but there may be a reluctance to share personal correspondence that contains sensitive information.

5.4 CONCLUSION

The impetus for conducting the suite of studies presented in this thesis was derived from the work by Lawlor and Kirakowski (2014). The inability to detect MbI or gauge the impact it may have on the dynamics of an OSG proved to be a limitation of their study. To remedy this, a suite of studies was conducted to develop a method to detect MbI and explore the dynamics of online communities when faced with MbI. While these aims were met, the findings had implications beyond the original impetus. The practical application value of the text classifier goes beyond a vetting tool for internet mediated researchers. It also can be used as an investigative tool and for moderators as a replacement for the unreliable high-level deception cues that are currently being used. The text classifier also had theoretical value. It supports the widely held theory that reduction of cognitive complexity is a feature of deception that can be identified through linguistic analysis. Furthermore, it supported the growing recognition that domain specificity must be included when interpreting LBC's of deception. To this end a linguistic profile of MbI was created. The development of the classifier also showed the benefits of using the SLP paradigm, which facilitated the inclusion of psychology theory when interpreting the cues, moving it beyond a simple classifier.

The groundwork exploratory studies of FD and MbI were used for choosing a method to detect MbI but also addressed the question of the dynamics involved in an online community faced with MbI and the aetiology of FD. The findings conflicted with traditional theories of FD obtained from observational studies and the new perspective of FD was used to suggest that FD be understood in terms of a behavioural addiction, which offers a more pragmatic understanding with implications for diagnosis, treatment and de-stigmatisation. The alignment of FD with addiction led to the conclusion that psychometric testing would not be a suitable method of detecting MbI

given their elusiveness and desire to remain hidden for fear of exposure. The most striking finding from the MbI study was that the presence or perceived presence of MbI limited opportunities to confer normality and form interpersonal support. MbI has the potential to erode the features of online communities that help to foster personal empowerment. MbI users were also found to be somewhat calculating in their deception by occupying 'ideal victim' personae. With regards to the development of a method to detect MbI, the common method used by online communities involved identifying observable cues embedded in the narratives created by MbI users. Given their ambiguity, more subtle cues which were not subject to human judgment bias were required and this led to the decision to develop a text classifier through machine learning.

The research conducted for this thesis not only addressed the issues which provided the original impetus for the studies, but given its exploratory nature it also had wider theoretical and practical application value. It has also provided a platform from which other studies can build on to deepen our understanding of FD and MbI beyond theoretical observations that have little pragmatic value and have stagnated research in this area. It requires a more inventive approach to acquiring data directly from those impacted by FD and MbI to broaden our understanding of the disorder and improve detection.

CHAPTER 6.0 CONSOLIDATED LIST OF REFERENCES

References

Adams, K.M. & Robinson, D.W. (2001). Shame reduction, affect regulation, and sexual boundary development: essential building blocks of sexual treatment addiction treatment. *Sexual Addiction and Compulsivity: The Journal of Treatment and Prevention 8,* 23–44.

Afroz, S., Brennan, M., & Greenstadt, R. (2011 May 20-23). *Detecting hoaxes, frauds, deception in online writing style*. Paper presented at 2012 IEEE Symposium on Security and Privacy: San Francisco. DOI 10.1109/SP.2012.34

Agarwal, C., & Mittal, N. (Eds.). (2016). *Socio-Affective Computing: Prominent Feature Extraction for Sentiment Analysis*. Switzerland: Springer.

Aggarwal, C.C., & Zhai, C.X. (2012). A survey of text clustering algorithms. In C.C. Aggarwal & C.X. Zhai (Eds.), *Mining Text Data* (pp.77-121). London: Springer.

Ahlberg, K., & Nordner, A. (2006). The importance of participation in support groups for women with ovarian cancer. *Oncology nursing form, 33*(4), 53-61.

Ahmedani, B.K. (2011). Mental health stigma: Society, individuals, and the profession. *Journal of social work values and ethics, 8*(2).

Akaichi, J. (2015). Sentiment classification: Facebook' statuses mining in the "Arabic Spring" era. In J. Zizka & F. Darena (Eds.), *Modern Computational Models of Semantic Discovery in Natural Language* (pp. 169-207). PA, USA: IGI Global.

Almela, A., Valencia-García, R., & Cantos, P. (2013). Seeing through Deception: A Computational Approach to Deceit Detection in Spanish Written Communication. *Linguistic Evidence in Security, Law and Intelligence, 1*(1). DOI: http://dx.doi.org/10.5195/lesli.2013.5

Alowibdi, J.S., Buy, U.A., & Yu, P. (2013 December 4-7). *Empirical Evaluation of Profile Characteristics for Gender Classification on Twitter*. Paper presented at The 12th International Conference on Machine Learning and Applications (ICMLA): Miami. DOI: 10.1109/ICMLA.2013.74

Amado, B.G., Arce, R., & Fariña, F. (2015). Undeutsch hypothesis and Criteria Based Content Analysis: A meta-analytic review European Journal of Psychology

Applied to Legal Context, 7, 3-12.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing

Artingstall, K. (1998). *Practical Aspects of Munchausen by Proxy and Munchausen Syndrome Investigation.* London: CRC Press.

Ayyer, S., & De Sousa. (2014). Cyber-Munchausen syndrome: A case report. *Indian Journal or Mental Health, 1*(1), 81-83.

Barak, A., & Boniel-Nissim, M., & Suler, J. (2008). Fostering empowerment in online support groups. *Computers in human behaviour, 24*(5), 1867-1883.

Bass, C., & Halligan, P.W. (2007). Illness related deception: Social or psychiatric problem? *Journal of The Royal Society of Medicine, 100*(2), 81-84.

Benlian, A., & Hess, T. (2011). The signalling role of IT features in influencing trust and participation in online communities. *International Journal of Electronic Commerce*, *15*(4).

Bhattacharya, K. (2017). *Fundamentals of Qualitative Research a Practical Guide.* Oxford, UK: Taylor and Francis.

Bird, S., Klein, E., & Looper, E., (2009). *Natural Language Processing with Python*. United States of America, CA: Sebastopol.

Birks, M., & Mills, J. (2015). *Grounded theory: A practical guide.* London, UK: SAGE.

Bishop, J. (2015). *Psychological and social implications surrounding internet and gaming addiction.* Pensylvania: Information science reference.

Black, D.W., Kehrberg, L.L.D., Flumerfelt, D.L. & Schlosser, S.S. (1997). Characteristics of 36 subjects reporting compulsive sexual behaviour. *American Journal of Psychiatry 154*, 243–249.

Bogdanova, D., Rosso, P., & Solorio, T. (2012, July). *On the Impact of Sentiment and Emotion Based Features in Detecting Online Sexual Predators*. In Proceedings of the 3rd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, Jeju, Republic of Korea, (pp. 110–118). Association for Computational Linguistics.

Böhm, A. (2004). Theoretical Coding: Text Analysis in Grounded Theory. In U. Flick, E. von Kardorff & I. Steinke (Eds.), *A Companion to Qualitative Research* (pp. 270-275). London, UK: SAGE Publications.

Bond, G.D., & Lee, A.Y. (2005). Language of lies in prison: linguistic classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology, 19*(3), pp. 313–329.

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214 –234.

Boyd, R. L., & Pennebaker, J.W. (2015). A way with words: Using language for psychological science in the modern era. In C.V. Dimofte, C.P. Haugtvedt & R.F. Yalch (Eds.), *Consumer Psychology in a Social Media World* (pp. 222-236). London, UK: Routledge.

Bowker, N.I., & Tuffin, K. (2006). Transcending operating barriers online for disabled bodies. *The Australian Journal of Rehabilitation Counselling, 12*(1), 46-61.

Braithwaite, D.O., Waldron., V.R., & Finn, J. (1999). Communication of social support in computer-mediated groups for people with disabilities. *Health Communication, 11*(2), 23- 51.

Brannon, G.E. (2015). *Factitious Disorder Imposed on Another*. Medscape. Retreived from: http://emedicine.medscape.com/article/295258-overview.

Breckenridge, J., Jones, D., Elliott, I., & Nicol, M. (2012) Choosing a methodological path: reflections on the constructivist turn. *Grounded Theory Review, 11*(1), pp. 64-71.

British Psychological Society (2013). *Ethics Guidelines for Internet-mediated Research.* Retrieved on October June 12th, 2015 from: http://www.bps.org.uk/system/files/Public%20files/inf206-guidelines-for-internet-mediated-research.pdf

Bryman, A. (2012). *Social Research Methods.* Oxford: Oxford University Press.

Buchanan, E.A. (2004). *Readings in virtual research ethics: Issues and controversies*. London: Information Science Publishing.

Buller, D. B., & Burgoon, J. K. (1994). Deception: Strategic and nonstrategic communication. In J. A. Daly & J. M. Weimann (Eds.), *Strategic interpersonal communication* (pp. 191-223). Hillsdale, NJ: Lawrence Erlbaum Associates.

Buller, D.B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203-242.

Buller, D. B., Burgoon, J. K., Buslig, A., & Roiger, J. (1996). Testing interpersonal deception theory: The language of interpersonal deception. *Communication Theory, 6,* 268–289.

Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2010). *Nonverbal Communication*. Boston: Allyn & Bacon.

Burns, M.B., & Moffitt, K.C. (2014). Automated deception detection of 911 call transcripts. *Security Informatics, 3*(8). DOI: 10.1186/s13388-014-0008-2

Cadet, P., & Feldman, M.D. (2012). Pretense of a paradox: Factitious intersex conditions on the internet. *International Journal of Sexual Health, 24*(2), 91-96.

Cagnina, L.C., & Rosso, P. (2015, September). *Tune or text: The effect of text on melody recognition*. In Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Lisbon, Portugal.

Carnes, P. (2001). Out of the Shadows: *Understanding Sexual Addiction,* 4th Ed. Hazelden, Minnesota.

Catalina, M.L., Gomez M.V., & de Cos, A. (2008). Prevalence of factitious disorder with psychological symptoms in hospitalized patients. *Actas Espanolas de Psiquiatria, 36*(6), 345-349.

Chamberlain, L. (2004). Understanding and diagnosing compulsive gambling. In: Coombs, R.H. (Ed.), *Handbook of Addictive Disorders: A Practical Guide to Diagnosis and Treatment.* John Wiley & Sons, Hoboken, New Jersey, pp. 126–158.

Charmaz, K. (1990). 'Discovering' chronic illness: using grounded theory. *Social Science and Medicine 30*, 1161–1172.

Charmaz, K. (1995). Grounded Theory. In J. A. Smith, R. Harré & L. van Langenhove (Eds.), *Rethinking Methods in Psychology* (pp. 27-49). London, UK: SAGE Publications.

Charmaz, K. (2003). Grounded theory. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods* (pp. 81–110). London: Sage

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Sage, London.

Charmaz, K. (2008). Constructionism and the grounded theory method. In J.A. Holstein & J.F. Gubrium (Eds.), *Handbook of constructionist research* (pp. 397–412). New York: The Guilford Press.

Charmaz, K. (2011). Grounded Theory Methods in Social Justice Research. In N.K. Denzin & Y.E. Lincoln (Eds.), *Handbook of Qualitative Research* (4th ed.). Lincoln: Sage.

Charmaz, K. (2014). *Constructing Grounded Theory* (2nd ed.). Sage: London.

Chen, D.L. (2012*). Does empathy beget guile?* Retrieved from: http://users.nber.org/~dlchen/papers/EmpathyGuile.pdf

Christie, N. (1986). The ideal victim. In E. Fattah (Ed.), *From crime policy to victim policy* (pp. 17-30), New York: St Martin's Press.

Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic indicators of psychological change after September 11, 2001. *Psychological Science, 15*, 687–693.

Conway, D., & White, J.M. (2012). *Machine Learning for Hackers.* O'Reilly, United States.

Cooney, A. (2010). Rigor and Grounded Theory. *Nurse Researcher, 14*(4), 17-22.

Corden, A., & Sainsbury, R. (2006). *Using verbatim quotations in reporting qualitative social research: researchers' views.* Social Policy Research Unit: University of York.

Corrigan, P., Markowitz, F.E., Watson, A., Rowan, D., & Kubiak, M.A. (2003). An attribution model of public discrimination towards persons with mental illness. *Journal of health and social behaviour*, *44*(2), 162-179.

Corrigan, P.W., Larson, J.E., & Rüsch, N. (2009). Self-stigma and the 'why try effect?': Impact on life goals and evidence based practices. *World Psychiatry, 8*(2), 75-81.

Corrigan, P.W., Rafacz, J., & Rusch, N. (2011). Examining a progressive model of self-stigma and its impact on people with serious mental illness. *Psychiatry Research*, 189(3), 339-342.

Coulson, N. S., Buchanan, H., & Aubeeluck, A. (2007). Social support in cyberspace: A content analysis of communication within a Huntington's disease online support group. *Patient Education and Counseling, 68*, 173-178.

Coulson, N.S. (2005). Receiving social support online: An analysis of a computer-mediated support group for individuals living with irritable bowel syndrome. *Cyberpsychology and behaviour, 8*(6), 580-584.

Coulson, N. S., & Greenwood, N. (2012). Families affected by childhood cancer: An analysis of the provision of social support within online support groups. *Child*

Care Health and Development, 38, 870-877.

Coursaris, K.C., & Liu, M. (2009). An analysis of social support exchanges in online HIV/AIDS self-help groups. *Computers in Human Behaviour, 25*(4), 911-918.

Criddle, L. (2010). Pediatric Care. Monsters in the Closet: Munchausen Syndrome by Proxy. *Critical Care Nurse*, *30*(6), 46-56.

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. London: Sage Publications Limited.

Cui, Y., & Zhao, Y. (2014). The Use of Second-Person Reference in Advertisement Translation with Reference to Translation between Chinese and English. *International Journal of Society, Culture and Language, 2*(1), 26-36.

Cunningham, J.M., & Feldman, M.D. (2011). Munchausen by Internet: current perspectives and three new cases. *Psychsomatics, 52*(2), 185-189.

Dadvar, M., de Jong, F., Ordelman, R., & Triesschnigg, D. (2012, February). *Improved Cyberbullying Detection.* In Proceedings of the Twelfth Dutch-Belgian Information Retrieval Workshop (DIR 2012), Ghent, Belgium (pp. 23-25). Belgium: Universiteit Gent.

Dayton, T. (2000). *Trauma and Addiction: Ending the Cycle of Pain through Emotional Literacy*. Health Communications, Florida.

Dearing, R.L., Stuewig, J. & Tangney, J.P. (2005). On the importance of distinguishing shame from guilt: relations to problematic alcohol and drug use. *Addictive Behaviours 30,* 1392–1404.

DePaulo, B.M. (1981). Success at Detecting Deception: Liability or Skill? *Annals of the New York Academy of Sciences, 364,* 245-255.

DePaulo, B.M., Malone, B.E., Lindsay, J.J., Muhlenbrook, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological bulletin, 129*(1), 74-118.

DiClemente, C.C., 2003. *Addiction and Change: How Addictions Develop and Addicted People Recover.* The Guilford Press, New York.

Dinakar, K., Reichart, R., & Lieberman, H. (2011). *Modeling the Detection of Textual Cyberbullying*. Paper presented at the Social Mobile Web Workshop at International Conference on Weblog and Social Media, Barcelona, Spain (pp. 11-17). Association for the Advancement of Artificial Intelligence.

Dingle, A.D. (2005). Disorders with physical symptoms. In S.B. Sexson (Ed.), *Child and adolescent psychiatry* (2nd ed.) (pp. 177-189). Oxford, UK: Backwell Publishing.

Donath, J.S. (1998). Identity and deception in the virtual community. In P. Kollock and M.A. Smith (Eds.), *Communities in Cyberspace* (pp. 27-59). London: Routledge.

Dryer, A.R., & Feldman, M.D. (2007). Factitious Disorder: Detection, Diagnosis, and Forensic Implications. *Psychiatric Times, 24*(4).

Dupre, E. (1909). Mythomanic infantile, Un cas de fugue suivie de fabulation. *Encephale, 11*.

Dzindolet, M. T., & Pierce, L. G. (2004, September). *A computerized text analysis can detect deception*. In Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, Santa Monica, CA. DOI: 10.1177/154193120404800377

Eastin, M.S., & LaRose, R. (2005). Alt.support: Modeling social support online. *Computers in human behaviour, 21*(6), 977-992.

Eastwood, S., & Bisson, J.L. (2008). Management of factitious disorders: A systematic review. *Psychotherapy & Psychosomatics, 77*(4), 209-218.

Eisendrath, S.J. (1996). Current overview of factitious physical disorders. In M.D. Feldman & S.J. Eisendrath (Eds.), *The spectrum of factitious disorders* (pp.21-35). Washington: American Psychiatric Press, Inc.

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing impressions online: Self-presentation processes in an online dating environment. *Journal of Computer Mediated Communication, 11*(2), 415-441.

Escalante, H.J., Villatoro-Tello, E., Juarez, A., Villasenor, L., & Montes-y-Gomez, M. (2013, June). *Sexual predator detection in chats with chained classifiers.* In Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Atlanta, Georgia, (pp. 46–54). Association for Computational Linguistics.

Esterly, R.W., & Neely, W.T. (1997). *Chemical Dependency and Compulsive Behaviours*. Lawrence Erlbaum Associates Inc., New Jersey.

Evans, G.L. (2013). A Novice Researcher's First Walk Through the Maze of Grounded Theory: Rationalization for Classical Grounded Theory. *The grounded theory review, 12*(1), 37-55.

Eysenbach, J.G. (2001). Ethical issues in qualitative research on internet communities. *British Medical Journal 323*, 1103–1105.

Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., & Stern, A. (2004). Health related virtual communities & electronic support groups: Systematic review of the

effects of online peer to peer interactions. *British Medical Journal, 328,* 1166-1171.

Eysenbach G., & Till J.E. (2001) Ethical issues in qualitative research on internet communities. *British Medical Journal, 323,* 1103–1105.

Feldman, M.D., & Eisendrath, S.J. (1996). *The Spectrum of Factitious Disorders.* American Psychiatric Press, Washington.

Feldman, M.D., Bibby, M., & Crites, S.D. (1998). 'Virtual' factitious disorders and Munchausen by proxy. *Western Journal of Medicine, 168*(6), 537-539.

Feldman, M.A., & Ala, B. (2000). Munchausen by Internet. *Southern medical journal, 93*(7).

Feldman, M.D. (2000). Munchausen by Internet: Detecting factitious illness and crisis on the internet. *Southern Medical Journal, 93*(7), 669-72.

Feldman, M.D., Hamilton, J.C., & Deemer, H.N. (2001). Factitious disorder. In: Philips, K. A. (Ed.), *Somatoform and Factitious Disorders*. American Psychiatric Publishing, Washington, pp. 129–159.

Feldman, M.D. (2004). *Playing sick? Untangling the web of munchausen syndrome, munchausen by proxy, malingering and factitious disorder.* UK: Brunner-Routledge.

Feldman, M.D., & Peychers, M.E. (2007). Legal issues surrounding the exposure of "Munchausen by Internet". *Psychosomatics, 48,* 451-2.

Fenner, R.M., & Gifford, M.H. (2014). Broadening the base of addiction mutual help oragnisations. In J.D. Roth, W.L. White & J.F. Kelly (Eds.), *Broadening the Base of Addiction Mutual Support Groups: Bringing Theory and science to contemporary trends* (pp. 4-25). London: Routledge.

Fielding, N., Lee, R.M., & Blank, G. (2008). *The SAGE handbook of online research methods.* London: SAGE.

Fire, M., Katz, G., & Elovici, Y. (2012). Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies. *Human Journal,* 1(1), 26–39.

First, M.B., & Tasman, A. (2010). *Clinical Guide to the Diagnosis and Treatment of Mental Disorders* (2nd ed.). Oxford, UK: John Wiley & Sons.

Fiske, S.T., Gilbert, D.T., & Lindzey, G. (2010). *Handbook of Social Psychology*. New Jersey: John Wiley & Sons.

Fleming, A.M., & Eisendrath, S.J. (2011). Factitious disorders. In A. Tasman, J. Kay, J.A. Liberman, M.B. First, & M. Maj (Eds.), *Psychiatry*. (3rd ed.) (UK: John Wiley and Sons.

Ford, C.V. (1996). *Lies! Lies!! Lies!! The Psychology of deceit*. Arlington: American Psychiatric Press.

Ford, C.V. (2010). Deception syndromes: factitious disorders and malingering. In: Levenson, J.L. (Ed.), *Textbook of Psychosomatic Medicine: Psychiatric Care of the Medically Ill* 2nd ed. American Psychiatric Publishing, Washington, pp. 291–303.

Fox, S., & Purcell, K. (2010 March 24). Chronic disease and the internet: Social Media and Health. *Pew Research Centre: Internet, science & tech.* Retrieved from: http://www.pewinternet.org/2010/03/24/social-media-and-health/

Frank, E., Hall, M.A., & Witten, I.H. (2016). *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*, Morgan Kaufmann, Fourth Edition, 2016.

Gaetano J. (2013). *Holm-Bonferroni sequential correction: An EXCEL calculator (1.2)* [Microsoft Excel workbook]. Retrieved from https://www.researchgate.net/publication/242331583_Holm-Bonferroni_Sequential_Correction_An_EXCEL_Calculator_-_Ver._1.2.

Gasson, S. (2009). Employing a Grounded Theory Approach for MIS Research. In Y.K. Dwivedi, B. Lal, M.D. Williams, S.L. Schneberger, & M. Wade (Eds.), *Handbook of research on Contemporary Theoretical Models in Information Systems* (pp. 34-56). Hersey PA: Information Science Reference.

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, USA: Aldine Publishing Company.

Glaser, B. G., & Strauss, A. L. (1968). *Time for dying*. Chicago, IL: Aldine.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. London, UK: Aldine Transaction.

Glaser, B. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sage.

Glaser, B. G. (1992). *Basics of grounded theory analysis.* Mill Valley, CA: Sociology Press.

Glaser, B. (1998). *Doing grounded theory: Issues and discussions*. Mill Valley, CA: Sociology Press.

Glaser, B. (2002). Constructivist grounded theory? *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 3*(3).

Glaser, B. (2005). *The grounded theory perspective 111: Theoretical coding.* Mill Valley, CA: Sociology Press.

Glaser, B. (2009). Jargonizing: The use of the grounded theory vocabulary. *The grounded theory review: An international journal 8*(1), 1-17.

Grazioli, S., and Jarvenpaa, S.L. 2003. Consumer and business deception on the Internet: Content analysis of documentary evidence. *International Journal of Electronic Commerce 7*(4), 93-118

Griffiths, M.K., Calear, L.A., Banfield, M., & Tam, A. (2009). Systematic review on internet support groups (ISGs) and depression (2): What is known about depression ISGs? *Journal of Medical Internet Research, 11*(3).

Grohol, J.M. (2006). *Anonymity and Online Community: Identity Matters*. Retrieved on from: http://alistapart.com/article/identitymatters.

Gudivada, V. N., Irfan, M.T, Fathi, T., & Rao, D.L. (2016). Cognitive analytics: Going beyond big data analytics and machine learning. In V. N. Gudvida, V.V. Raghavan, V. Govindaraju & C.R. Rao (Eds.), *Cognitive computing: Theory and applications* (pp. 169-207). Oxford, UK: Elseiver.

Gupta, S. & Skillicorn, D. (2006, January). *Improving a Textual Deception Detection Model.* In Proceedings of the 2006 Conference of the Center for Advanced Studies on Collaborative Research, Toronto, Canada. DOI: 10.1145/1188966.1189005

Gupta, A., Kumaraguru, P., & Sureka, A. (2012). *Characterizing Paedophile Conversations on the Internet using Online Grooming*. Computing Research Depository. arXiv:1208.4324.

Gustavsson, B. (2007). Grounded theory. In B. Gustavsson (Ed.), *The Principles of Knowledge Creation: Research Methods in the Social Sciences* (pp. 66-79). Cheltenham, UK: Edward Elgar.

Guzman, J. & Correll, T. (2008). Factitious disorder. *Psychiatry 11*, 3–10.

Hagglund, L.A. (2009). Challenges in the treatment of factitious disorder: a case study. *Archives of Psychiatric Nursing, 23*, 58–64.

Hall, L. & Lloyd, S. (1993). *Surviving child sexual abuse*. London: The falmer press.

Hall, M., Fran, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I.H. (2009). The WEKA Data Mining Software: An Update. *SIGKDD Explorations, 11*(1).

Hall, P. (2012). *Understanding and Treating Sex Addiction.* Routledge, New York.

Hall, J.M. (2013). *Simple Solution to the Troll / Hacker / Cyberbullying Problem.* Retrieved: http://getsatisfaction.com/ imdb/topics/simple_ solution_to_the_ troll_hacker_cyberbullying_problem.

Hallstrom, C. (2009). *Munchausen Syndrome (Factitious Disorder).* Good Medicine Practice.

Hamilton, J.C., & Janata, J.W. (1997). Dying to be ill: the role of self-enhancement motives in the spectrum of factitious disorders. *Journal of Social and Clinical Psychology 16*, 178–199.

Hamilton, J.C., Feldman, M.D., & Janata, J.W. (2009). The A, B, C's of factitious disorder: A response to Turner. *The Medscape Journal of Medicine, 11*(1).

Hancock, J.T., Curry, L.E., Goorha, S., & Woodworth, M. (2008). On Lying and Being Lied To: A Linguistic Analysis of Deception in Computer-Mediated Communication, *Discourse Processes, 45*(1), 1-23.

Hancock, J.T., Beaver, D.I., Chung, C.K., Frazee, J., Pennebaker, J.W., Graesser, A., & Cai, Z. (2010). Social language processing: A framework for analyzing the communication of terrorists and authoritarian regimes. *Behavioral Sciences of Terrorism and Political Aggression, 2*(2), 108-132.

Hancock, J. T. (2011). Digital deception. Why, when and how people lie online. In A. Joinson, K. McKenna, T. Postmes & U. Reips (Eds.), *The Oxford Handbook of Internet Psychology* (pp. 289-301). Oxford: Oxford University Press.

Hartwig, M., & Bond, C.F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological bulletin, 137*(4), 643-659.

Hauch, V., Blandon-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are computers effective lie detectors? *A meta-analysis of linguistic cues to deception, 19*(4), 307-42.

Helgeson, V.S., & Gottlieb, B.H. (2000). Support groups. In S.Cohen, L.G. Underwood & B.H. Gottlieb (Eds,). *Social Support Measurement and Intervention: A Guide for Health and Social Scientists* (pp. 221-246). UK: Oxford University Press.

Hernandez Fusilier, D., Montes-y-Gomez, M., Rosso, P., & Guzman Cabrera, R. (2015). Detection of ´ opinion spam with character n-grams. In A. Gelbukh (Ed.), *Lecture Notes in Computer Science: Computational Linguistics and Intelligent Text Processing* (pp. 285–294). New York: Springer.

Herring, S., Job-sluder, K., Scheckler, R., & Barab, S. (2002). Searching for Safety Online: Managing "Trolling" in a Feminist Forum. *The information society*, *18*, 371-384.

Hewson, C. (2003). *Internet research methods: A practical guide for the social and behavioural sciences*. London: SAGE.

Higginbottom G, & Lauridsen E.I. (2014). The roots and development of constructivist grounded theory. *Nurse Researcher, 21*(5), 8-13.

Hirschberg, J., Benus, S., Brenier, J.M., Enos, F., Friedman, S., Gilman, S., Girand, C., Graciarena, M., Kathol, A., Michaelis, L., Pellom, B., Schriberg, E., & Stolcke, A. (2005). *Distinguishing deceptive from non-deceptive speech.* In Proceedings of Interspeech 2005, Lisbon: Spain. Retrieved from: http://www.cs.columbia.edu/~julia/files/hirschberg_al_05.pdf

Hoijer, B. (2004). The discourse of global compassion: The audience and media reporting of human suffering. *Media, culture and society*, 26(4), 513-531.

Holton, J. A. (2010). The Coding Process and Its Challenges. *The Grounded Theory Review, 9*(1), 21-41.

Hwang, K.O., Ottenbacher, A.J., Green, A.P., Cannon-Diehl, M.R., Richardson, O., Bernstam, E.V., & Thomas, E.J. (2010). Social support in an Internet weight loss community. *International journal of medical informatics*, *79*(1), 5-13.

Idriss, S.Z., Kvedar, J.C., & Watson, A.J. (2009). The role of online support communities: Benefits of expanded social networks to patients with psoriasis. *Archives of dermatology*, *145*(1), 46-51.

Ishak, W.W., Rasyidi, E., Saah, T., Vasa, M., Ettekal, A., & Fan, A. (2010). Factitious disorder: Case series with variations of psychological and physical symptoms. *Primary psychiatry*, *17*(9), 40-43.

Jarvis, S. (2012). The detection-based approach: An overview. In S. Jarvis & S.A. Crossley (Eds.), *Approach language transfer through text classification* (pp. 1-34). London, UK: Multilingual Matters.

Jenkins, J. (2012 June 01). The cruelest of internet hoaxes. *BBC News.* Retrieved from: http://www.bbc.co.uk/news/magazine-18282277.

Jensen, E. & Laurie, C. (2016). Doing Real Research, A Practical Guide to Social Research. London, UK: Sage.

Jeter, P. (2015). *The stranger among us: Identity deception in online communities of choice*. Retrieved. http://www.academia.edu /1629343/ The_Stranger_ Among_Us_Identity_ Deception_in_Online_Communities_of_Choice

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67- 85.

Joinson, A.N. (2001). Self-disclosure in computer mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology, 31*(2), 177-192.

Joost, H & Schulman, A. (Directors). (2010). *Catfish* [Documentary]. United States of America: Relativity Media.

Jordon, K.N., & Pennebaker, J.W. (2016). How the Candidates are Thinking: Analytic versus Narrative Thinking Styles. Retrieved from: https://wordwatchers.wordpress.com/2016/01/21/how-the-candidates-are-thinking-analytic-versus-narrative-thinking-styles/

Kacewicz, W., Pennebaker, J.W., Davis, M., Jeon, M., & Graesser, A.C. (2013). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*. DOI: 10.1177/0261927X1350265.

Kahn, S.A. (2008). Munchausen syndrome by proxy. *Physician's academy, 2*(5), 5- 7.

Kanaan, R.A., & Wessely, S.C. (2010). Factitious disorders in neurology: An analysis of reported cases. *Psychosomatics, 51*(1), 47-54.

Kanaris, I., Kanaris, K., & Samatatos, E. (2007). *Spam detection using character n-grams.* In G. Antoniou, G. Potamias, C. Spyropoulos & D. Plexousakis (Eds.), In Proceedings of the 4th Helenic conference on Advances in Artificial Intelligence (pp. 95-104). Berlin: Springer.

Kaplan, A.M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons, 53,* 59-68.

Karami, A., & Zhou, B. (2015). *Online Review Spam Detection by New Linguistic Features*. Paper presented at the iConference 2015 Proceedings, Illinois, United States.

Kelle, U. (2005). "Emergence" vs. "forcing" of empirical data? A crucial problem of "grounded theory" reconsidered. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 6*(2), 1270-1289.

Kendall, J. (1999). Axial coding and the grounded theory controversy. *Western Journal of Nursing Research, 21*(6), 743-757.

Kenny, M., & Fourie, R. (2015). Contrasting Classic, Straussian, and Constructivist Grounded Theory: Methodological and Philosophical Conflicts. *The Qualitative Report, 20*(8).

Krahn, L.E., Li, H., & O'Connor, K.M. (2003). Patients who strive to be ill: Factitious disorder with physical symptoms. *American Journal of Psychiatry, 160*(6), 1163-1168.

Krahn, L.E., Bostwick, J.M. & Stonnington, C.M. (2008). Looking toward DSM-V: should factitious disorder become a subtype of somatoform disorder? *Psychosomatics 49*, 277–282.

Kiesler, S., Kittur, A. Kraut, R., Resnick, P. (2010). Regulating behaviour in online communities. In R.E. Kraut & P. Resnick (Eds.), *Evidence based social design: Mining the social sciences to build online communities* (pp. 125-179). Massachusetts: MIT Press.

King, B.H., & Ford, C.V. (1988). Pseudologia fantastica. *Acta Psychiatra Scandinavica, 77*(1), 1-6.

Kim, K., Ryu, E., Chon, M.Y., Yeun, E.J., Choi, S.Y., Seo, J.S. & Nam, B.W. (2006). Internet addiction in Korean adolescents and its relation to depression and suicidal ideation. A questionnaire survey. *International Journal of Nursing Studies 43*, 185–192.

Kleeman, J. (2011 February 26). Sick note: Faking illness online. *The Guardian.* Retrieved from: http://www.theguardian.com/lifeandstyle/2011/feb/26/faking-illness-online-munchausen.

Klein, R., & Dinger, M. (2008). Trust in patient-centred E-Health. In E.V. Vance (Ed.), *Patient centered E-Health* (pp.98-114). New York: Medical Information Science Reference.

Köhnken, G. (2004). Statement Validity Analysis and the detection of the truth. In P. A. Granhag & L. A. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 41-63). Cambridge, UK: Cambridge University Press.

Kotsiantis, S.B. (2011). Feature selection for machine learning classification problems: a recent overview. *Artificial Intelligence Review, 42*(2).

Kuiken, D. (1981). Nonimmediate language style and inconsistency between private and expressed evaluations. *Journal of Experimental Social Psychology, 17*, 183-196.

Kumar, S., Cheng, J., Leskovec, J., & Subrahmanian, V.S. (2017, April). *An Army of*

Me: Sockpuppets in Online Discussion Communities. Paper presented at 26th International World Wide Web conference, Perth, Australia. arXiv:1703.07355

Kummervold, P.E., Gammon, D., Bergvik, S., Johnsen, J.A., Hasvold, T., & Rosenvinge, J.H. (2002). Social support in a wired world: use of online mental health forums in Norway. *Nordic Journal of Psychiatry*, 56(1), 59-65.

Kuntsman, A. (2007). Belonging through violence: Flaming, erasure, and performativity in queer migrant community. In K. O'Riordan & D.J. Phillips (Eds.), *Queer Online: Media technology and sexuality* (pp. 101-123). New York: Peter Lang Publishing.

Kurtz, L.F. (2015). *Recovery Groups: A guide to creating, leading, and working with groups for addictions and mental health conditions*. New York: Oxford University Press.

Kuruvilla, A.M., & Varghese, S. (2015, March). *A detection system to counter identity deception in social media applications*. Paper presented at the International Conference on Circuit, Power and Computing Technologies (ICCPCT), Nagercoil, India. IEEE: DOI: 10.1109/ICCPCT.2015.7159321.

Lantz, B. (2015). *Machine Learning with R*. Birmingham, UK: Packt Publishing Ltd.

Lasher, L.J., & Sheridan, M.S. (2004). *Munchausen by Proxy: Identification, intervention, and case management.* New York, NY: Routledge.

Lawlor, A., & Kirakowski, J. (2014). Online support groups for mental health: A space for challenging self-stigma or a means of social avoidance? *Computers in human behaviour, 32,* 152-162.

Lawlor, A. & Kirakowski, J. (2016). *Claiming someone else's pain: A Grounded theory analysis of online community user's experiences of Munchausen by Internet.* Manuscript under review.

LeCroy, C.W., & Holschuh, J. (2012). *First person accounts of mental illness and recovery*. New Jersey, US: John Wiley & Sons.

Leins, D.A., Fisher, R.P., & Vrij, A. (2012). Drawing on Liars' Lack of Cognitive Flexibility: Detecting Deception Through Varying Report Modes. *Applied Cognitive Psychology, 26*(4), 601-607.

Levine, T.R., & Kim, R.K. (2010). Some considerations for a new theory of deceptive communications. In M.S. McGlone & M.L. Knapp (Eds.), *The Interplay of Truth and Deception: New Agendas in Theory and Research* (pp. 16-35). Oxon, UK: Routledge.

Levitt, H.M. (2015). Interpretation-driven guidelines for designing and evaluating grounded theory research: A constructivist social justice approach. In O.C.G. Gelo., A. Pritz., & B. Rieken (Eds.), *Psychotherapy Research: Foundations, Process and Outcome* (pp. 455-485). London, UK: Springer.

Li, J., Cardie, M. O. C., and Hovy, E. (June, 2014). *Towards a general rule for identifying deceptive opinion spam.* In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, Maryland, USA (pp. 1566-1576). Association for Computational Linguistics.

Lloyd, A. (2007). Understanding information literacy in the workplace: Using a constructivist grounded theory approach. In S. Lipu., K. Williamson., & A. Lloyd (Eds.), *Exploring Methods in Information Literacy Research* (pp. 68-86). Wagga Wagga: Centre for Information Studies, Charles Sturt University.

Loewenstein, G., & Small, D.A. (2007). The Scarecrow and the Tin Man: The vicissitudes of human sympathy and caring. *Review of general psychology, 11*(2), 112-126.

Long, T., & Johnson, M. (2000). Rigour, reliability and validity in qualitative research. *Clinical Effectiveness in Nursing, 4*(1), 30-37.

Long, W.R. (2008). *Munchausen Syndrome By Proxy/Factitious Disorder By Proxy*. Autism Research Institute. Retrieved from: http://www.autism.com /pdf/fami lies / MBSP_unabridged.pdf

MacCarthy, M., Kupiers, L., Hurry, J., Harper, R., & LeSage, A. (1989). Counselling the relatives of the long-term adult mentally ill. I. Evaluation of the impact on relatives and patients. *British Journal of Psychiatry, 154,* 768-75.

Maldonado, J.R. (2002). When patients deceive doctors: A review of factitious disorder. *American journal of forensic psychiatry*, *23*(2), 30-58.

Markham, A.N. (2012). Fabrication as ethical practice: Qualitative inquiry in ambiguous internet contexts. *Information, Communication & Society, 3,* 334-353.

Masip, J., Sporer, S.L., Garrido, E, & Herrero, C. (2005). The detection of deception with the Reality Monitoring approach: A review of the empirical evidence. *Psychology, crime, and law, 11*(1), 99-122.

McClure, R.J., Davis, P.M., Meadow, S.R., & Sibert, J.R. (1996). Epidemiology of Munchausen syndrome by proxy, non-accidental poisoning, and non-accidental suffocation. *Archives of disease in childhood, 75*(1), 57-61.

McCord, M., & Chuah, M. (2011). Spam detection on Twitter using traditional classifiers. In J.M.A. Calero, L.T. Yang, F.G. Marmol, L.J.G. Villalba, A.X. Li, & Y. Wang (Eds.), *Autonomic and Trusted Computing* (pp. 175-186). Berlin: Springer.

McCormack, A., & Coulson, S.N. (2005). Individuals with eating disorders and the use of online support groups as a form of social support. Cyberpsychology: *Journal of Psychosocial Research on Cyberspace*, *3*(2), 1-12.

McCulloch, V., & Feldman, M.D. (2011). Munchausen by proxy by Internet. *Child abuse & neglect*, *35*(11), 965-6.

McDermott, B.E., Leamon, M.H, Feldman, M.D. & Scott, C.L. (2000). Factitious disorder and malingering. In: Hales, R.E., Yudofsky, S.C., Gobbard, G.O. (Eds.), *Textbook of Psychiatry*, 5th ed. American Psychiatric Publishing, Virginia, pp. 643–664.

McIntosh, J. & McKeganey, N.P. (2002). *Beating the Dragon: the Recovery from Dependent Drug Use*. Pearson Education, London.

McKee, H.A., & Porter, J.E. (2009). *The ethics of internet research*: A rhetorical case based

approach. New York: Peter Lang Publishing.

McKenna, K.Y.A. (2008). Influence on the nature and functioning of online groups. In A. Barak (Ed). *Psychological aspects of cyberspace: Theory, research and application*. UK: Cambridge University Press.

McLaughlin, B., Hull, S., Namkoong, K., Shah, D., & Gustafson, D.H. (2016). We all scream for ice cream: Positive identity negotion in the face of cancer. In A. Novak & I.J. El-Burki (Eds.), *Defining Identity and the Changing Scope of Culture in the Digital Age* (pp. 81-99). Penslyvania: IGI Global.

Meadows, S.R. (1982). Munchausen's syndrome: when a doctor turns detective. *Nursing Mirror 154,* 17–20.

Miah, W.R., Yearwood, J., & Kulkarni, S. (2011). *Detection of child exploiting chats from a mixed chat dataset as a text classification task*. In Proceedings of Australasian Language Technology Association Workshop, Caulfield, Melbourne, Australia (pp. 157−165). Association for Computational Linguistics.

Mihalcea, R., & Strapparava, C. (2009, august). *The lie detector: Explorations in the automatic recognition of deceptive language.* In Proceedings of the ACL-

IJCNLP 2009 Conference Short Papers, Suntec, Singapore.

Miller, R.D. (2007). Forensic aspects of factitious disorder. In L.B. Schlesinger (Ed,). *Explorations in criminal psychopathology: Clinical syndromes with forensic implications* (2nd ed) (pp. 158-186). Illinois: Charles C Thomas, Publisher Ltd.

Mills, J., Bonner, A., & Francis, K. (2006). Adopting a constructivist approach to grounded theory: implications for research design. *International Journal of Nursing Practice, 12*(1), 8-13.

Miyata, K. (2008). Social Support for Japanese mothers online and offline. In B. Wellman & C. Haythornthwaite (Ed.), *The Internet in Everyday Life* (pp. 520-549). UK: Blackwell Publishing.

Mo, P.K., & Coulson, N.S. (2008). Exploring the communication of social support within virtual communities: A content analysis of message posted to an online HIV/AIDS support group. *Cyberpsychology and Behaviour, 11*, 371-374

Mo, P.K., & Coulson, N.S. (2014). Are online support groups always beneficial? A qualitative exploration of the empowering and disempowering processes of participation within HIV/AIDS-related online support groups. *International Journal of Nursing Studies, 51*, 983-93.

Monroe, R. (2016 February 18). *Cancer cons, phoney accidents and fake deaths: meet the internet hoax buster. The Guardian*. Retrieved from: https://www.theguardian.com/technology/2016/feb/18/cancer-cons-phoney-accidents-fake-deaths-internet-hoax-buster-taryn-wright.

Montague, J. (2015 April 29). Munchausen by Internet: The sickness bloggers who fake it online. *The Guardian*. Retrieved on the 12th September, 2015 from: http://www.theguardian.com/society/2015/apr/29/jules-gibson-munchausen-by-internet-sickness-bloggers-fake-it-whole-pantry.

Moran, J.M. (2012 October 08). Faking it online: The disturbing trend of Internet cancer hoaxes. *Fox News Health*. Retrieved on the 2nd of April, 2012 from: http://www.foxnews.com/health/2012/10/08/faking-it-online-disturbing-trend-internet-cancer-hoaxes.html

Morse, J.M., Stern, P.N., Corbin, J., Bowers, B., Charmaz, K., & Clarke, A.E. (2016). *Developing Grounded Theory: The Second Generation.* London, UK: Routledge.

Murali, V., Ray, R., & Shaffiullha, M. (2012). Shopping addiction. *Advances in Psychiatric Treatment 18*, 278–279.

Munhall, P. (2007). The landscape of qualitative research. In P. Munhall (Ed.), *Nursing Research: A Qualitative Perspective* (4th ed, pp. 3-36). Sudbury, MA: Jones and Bartlett.

National Research Council, Committee on the Social and Economic Impact of Pathological Gambling (1999). *Pathological Gambling: A Critical Review*. National Academies Press, Washington.

Newman, M.L., Pennebaker, J.W., Berry, D.S., & Richards, J.M. (2003). Lying Words: Predicting Deception From Linguistic Styles. *Personality and Social Psychology Bulletin, 29,* 665-675.

Nimrod, G. (2012a). The Membership Life Cycle in Online Support Groups. *International Journal of Communication, 6,* 1245–126.

Nimrod, G. (2012b). Online Depression Communities: Does Gender Matter? *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 6*(1).

Nimrod, G. (2013). Probing the audience of seniors' online communities. *The Journals of gerontology. Series B, Psychological sciences and social sciences*, *68*(5), 773-82.

Obst, P., & Stafurik, J. (2010). Online we are all able bodied: Online psychological sense of community and social support found through membership of disability-specific websites promotes well-being for people living with a physical disability. *Journal of Community & Applied Social Psychology*, *20*(6), 525-531.

Orford, J. (2001). Conceptualising addiction: addiction as excessive appetite. *Addiction 96,* 15–31.

O'Shea, B. (2003). Factitious disorder: the Baron's legacy. *International Journal of Psychiatry in Clinical Practice 7*, 33–39.

Ott, M., Choi, Y., Cardie, C., & Hancock, J. (2011). *Finding deceptive opinion spam by any stretch of the imagination*. Paper presented at 49th Annual Meeting of the Association for Computational Linguistics, Portland, Oregon.

Ott, M., Cardie, C., & Hancock, J. (2012). *Estimating the prevalence of deception in online review communities.* Paper presented at the World Wide Web Conference Commitee, Lyon, France.

Palmer, R. (2013, January 17). *Where does 'Catfish' come from? Online hoax movie inspired by fisherman's lore*. International Business Times. Retrieved from: http://www.ibtimes.com/where-does-catfish-come-online-hoax-movie-inspired-fishermans-lore-1022374

Pankratz, L. (2006). Persistent problems with the Munchausen syndrome by proxy label. *Journal of the American Academy of Psychiatry and the Law Online, 34*(1), 90-5.

Pavia, W. (2015 March 4). *Mother who blogged about sick son, 5, was killing him with salt*. *TheTimes.* Retrieved from: http://www.thetimes.co.uk/tto/news/world/americas/article4371173.ece

Peersman, C., Daelemans, W., & Van Vaerenbergh, L. (2011, October). *Predicting Age and Gender in Online Social Networks*. In Proceedings of the 3rd international workshop on Search and mining user-generated contents, Glasgow, Scotland, UK, (pp. 37-44). New York, USA: Association of Computing Machinery.

Pendar, N. (2007, September). *Toward spotting the pedophile: Telling victim from predator in text chats*. In Proceedings of the International Conference on Semantic Computing, Irvine, California, (pp. 235–241). IEEE. DOI: 10.1109/ICSC.2007.32

Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: University of Texas.

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). The development and psychometric properties of LIWC2007 [LIWC manual]. Austin, TX: LIWC.net.

Perlmutter, R.A. (2004). *A family approach to psychiatric disorders*. UK: American psychiatric press, Inc.

Phoenix, K.H.M., & Coulson, N.S. (2010). Empowering processes in online support groups among people living with HIV/AIDS: A comparative analysis of 'lurkers' and 'posters'. *Computers in Human Behaviour, 26*(5), 1183-1193.

Preece, J. (1999) Empathic communities: Balancing emotional and factual communication. *Interacting with Computers, 12*, 63-77.

Preece, J., & Maloney-Krichmar, D. (2003). Online communities: Focusing on sociability and usability. In J. Jacko & A. Sears (Eds.), *Handbook of Human Computer Interaction* (pp. 596-620). New Jersey: Lawrence Erlbaum Associates Inc.

Pridmore, S. (2006). *Psychiatry*. Tasmania: University of Tasmania.

Pelletiere, N. (2015 January 9). *Photos of Sick Child Stolen, Posted to Fake Instagram Account. Abc News*. Retrieved from: http://abcnews.go.com/Lifestyle/photos-sick-child-stolen-posted-fake-instagram-account/story?id=28094385

Pulman, A., & Taylor, J. (2012). Munchausen by Internet: Current Research and Future Directions. *Journal of Medical Internet Research, 14*(4).

Qin, T. T., Burgoon, J., Blair, J. P., & Nunamaker Jr, J. F. (2005). *Modality effects in deception detection and applications in automatic-deception-detection.* In Proceedings of the 38th Hawaii International Conference on System Sciences, Hawaii, US.

Rains, S. A., & Young, V. (2009). A meta-analysis of research on formal computer-mediated support groups: Examining group characteristics and health outcomes. *Human Communication Research, 35*(3), *309-336.*

Ramalho, R., Adams, P., Huggard, P., & Hoare, K. (2015). Literature Review and Constructivist Grounded Theory Methodology. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 16*(3).

Raschka, S. (2014). *Naive Bayes and Text Classification I Introduction and Theory*. eprint arXiv:1410.5329.

Reich, P., & Gottfried, L.A. (1983). Factitious disorders in a teaching hospital. *Annals of internal medicine, 99*(2), 240-247.

Rice, D.R., & Zorn, C. (2013, September). *Corpus-Based Dictionaries for Sentiment Analysis of Specialized Vocabularies*. Paper presented at the New Directions in Analyzing Text as Data Workshop, London, UK.

Ridings, C.M., Gefen, D., & Arinze, B. (2002). Some antecedents and effects of trust in virtual communities. *Journal of Strategic Information Systems, 11*(3-4), 271-295.

Ries, R.K., Fiellin, D.A., Miller, S.C. & Saitz, R. (2009). *Addiction Medicine,* 4th ed. Lippincott, Williams & Wilkins, Philadelphia.

Rodgers, S., & Chen, Q. (2006). Internet Community Group Participation: Psychosocial Benefits for Women with Breast Cancer. *Journal of Computer Mediated Communication, 10*(4).

Rowe, N.C. (2005). Exploiting the captions for multimedia data mining. In P. Margherita (Ed,). *Encyclopaedia of multimedia technology and networking* (pp. 302-308), London: Idea Group Reference.

Russell, B., & Purcell, J. (2009). *Online research essentials*: Designing and implementing Research studies. San Francisco: John Wiley & Sons.

Sadock, B.J., & Sadock, V.A. (2008). *Kaplan and Sadock's concise textbook of clinical psychiatry* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Sadock, B.J., & Sadock, V.A. (2011). *Kaplan and Sadock's synopsis of psychiatry* (10th ed.). Pennsylvania: Lippincott Williams & Wilkins.

Saunders, B., Kitzinger, J., & Kitzinger, C. (2014). Participant anonymity in the internet age: from theory to practice. *Qualitative Research in Psychology, 12*(2), 125-137.

Sanderson, C. (2006). *Counselling adult survivors of child sexual abuse* (3rd ed). London: Jessica Kingsley Publishers.

Sargeant, J. (2012). Qualitative Research Part II: Participants, Analysis, and Quality Assurance. *Journal of graduate medical education, 4*(1), 1-3.

Sarkar, D. (2016). *Text analytics with python: A practical real-world approach to gaining actionable insights from your data*. New York: Apress.

Savino, A.C., Fordtran, J.S. (2006). Factitious disease: clinical lessons from case studies at Baylor University Medical Center. *Baylor University Medical Centre Proceedings, 19*(3), 195-208.

Sbaraini, A, Carter, S.M., Evans, R.W., & Blinkhorn, A. (2011). How to do a grounded theory study: a worked example of a study of dental practices. *BMC Medical Research Methodology, 11*(128).

Schelleman-Offermans, K., & Merckelbach, H. (2010). Fantasy Proneness as a Confounder of Verbal Lie Detection Tools. *Journal of Investigative Psychology and Offender Profiling, 7*, 247-260.

Segovia, K.Y. & Bailenson, J.N. (2013). Identity Manipulation: What Happens When Identity Presentation is Not Truthful. In Amichai-Hamburger Y. (Ed.), *The Social Net: Understanding Our Online Behaviour* (pp.45-61). UK: Oxford University Press.

Sengstok, B. (2015). A Grounded Theory Study of Nursing Students' Experiences in the Off-Campus Clinical Setting. In De Chesnay, M (Ed.), *Nursing Research Using Grounded Theory* (pp.169-203). Springer: London.

Sergienko, R., Shan, M., & Schmitt, A. (2016). A comparative study of text preprocessing techniques for natural language call routing. In K. Jokinen & G.

Wilcock (Eds.), *Dialogues with Social Robots: Enablements, Analyses, and Evaluation* (23-39). Springer: London.

Setoyama, Y., Yamazaki, Y., & Namayama, K. (2011). Benefits of peer support in online Japanese breast cancer communities: Differences between lurkers and posters. *Journal of Medical Internet Research*, *13*(4).

Schreier, H.A., & Libow, J.A. (1993). Munchausen syndrome by proxy: Diagnosis and prevalence. *American Journal of Orthopsychiatry*, *63*(2), 318-321.

Sheridan, M.S. (2003). The deceit continues: an updated literature review of Munchausen Syndrome by Proxy. *Child abuse & neglect, 27*(4), 331-451.

Slutske, W.S. (2006). Natural recovery and treatment-seeking in pathological gambling: results of two U.S. national surveys. *American Journal of Psychiatry 163*, 297–302.

Sobel, L.C. (2007). The phenomenon of self-change: overview of key issues. In: Klingemann, H., Sobell, L.C. (Eds.), *Promoting Self-change from Addictive Behaviours*. Springer, New York, pp. 1–31.

Solorio, T., Hasan, R., & Mizan, M. (2013a, June). *A Case Study of Sockpuppet Detection in Wikipedia.* In Proceedings of the Workshop on Language in Social Media (LASM 2013), Atlanta, Georgia, USA, (pp. 59-68). Association for Computational Linguistics.

Solorio, T., Hasan, R., & Mizan, M. (2013b). *Sockpuppet detection in Wikipedia: A corpus of real-world deceptive writing for linking identities*. Computing Research Repository. arXiv:1310.6772.

Somekh, B., & Lewin, C. (2005). *Research methods in the social sciences.* London: SAGE.

Sporer, S. L., Kupper, B. and Bursch, S. E. (April, 1995). *Hilft Wissen uber Realitatsu berwachung, um zwischen wahren und erfundenen Geschichten zu unterscheiden?* [Does knowledge about reality monitoring help to discriminate between true and invented stories?]. Paper presented at the 37th Tagung experimentell arbeitender Psychologen, Bochum, Germany.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Newbury Park, CA: Sage

Steel, R.M. (2009). *Factitious disorder (Munchausen's syndrome*). The Journal of the Royal College of Physicians Edinburgh 39, 343–347.

Steller, M., & Köhnken, G. (1989). Criteria-Based Content Analysis. In D. C. Raskin

(Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217-245). New York, NY: Springer-Verlag.

Stommel, W. (2009). *Entering an Online Support Group on Eating Disorders: A Discourse Analysis*. New York: Rudopi.

Swains, H. (2009, March 25). *Q&A: Munchausen by internet* [Blog Post]. Retrieved from: http://www.wired.co.uk/news/archive/2009-03/24/qa-munchausen-by-internet.aspx

Szalavitz, M. (2012 November 30). Munchausen By Internet: Faking Illnesses Online. *Time.* Retrieved from: http://healthland.time.com/2012/11/30/munchausen-by-internet-faking-illnesses-online/

Tasman, A., Mohr, W.K. (2011). *Fundamentals of Psychiatry*. John Wiley & Sons, Oxford.

Tausczik, Y.R. & Pennebaker, J.W. (2010). The psychological meaning of words. LIWC and computerized text analysis methods. *Journal of language and social psychology, 29(*1), 24-54.

Taylor, S. & Hyler, S.E. (1993). Update on factitious disorder. *International Journal of Psychiatry in Medicine 23*, 81–94.

Toma, C. L., & Hancock, J. T. (2010). Looks and lies: Self–presentation in online dating profiles. *Communication Research, 37*, 335–531.

Toma, C. L., & Hancock, J. T. (2012). What Lies Beneath: The Linguistic Traces of Deception in Online Dating Profiles, *Journal of Communication*, 62(1), 78-97.

Toneatto, T. (2013). Natural recovery. In: Miller, P.M. (Ed.), *Principles of Addiction: Comprehensive Addictive Behaviours and Disorders.* Academic Press, London, pp. 133–141.

Tonkin, E. L. (2016). A day at work (with text): A brief introduction. In E.L. Tonkin & G.J.L. Tourte (Eds.), *Working with Text: Tools, Techniques and Approaches for Text Mining* (pp. 1-19). Oxford, UK: Elseiver.

Trondsen, M., & Tjora, A. (2015). Communal Normalization in an Online Self-Help Group for Adolescents With a Mentally Ill Parent. *Qualitative health research*, *24*(10), 1407-17.

Turner, M.A. (2006). Factitious disorders: reformulating the DSM-IV criteria. *Psychosomatics 47*, 23–32.

Tsikerdekis, M., & Zeadally, S. (2014). Online Deception in Social Media. *Communications of the ACM, 57*(9), 72-80.

Undeutsch, U. (1989). The development of statement reality analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 101–121). Dordrecht, Netherlands: Kluwer Academic Publishers.

Uridge, L., Rodan, D., & Green, L. (2012). *A Moderator's Dilemma: Munchausen Syndrome By Internet.* Paper presented at the ANZCA, Adelaide, South Australia.

van de Loo, J., De Pauw G., & Daelemans W. (2016). Text-Based Age and Gender Prediction for Online Safety Monitoring. *The International Journal of Cyber-Security and Digital Forensics, 5*(1), 46-60.

Van Uden-Kraan, C.F., Drossaert, C.H.C., Taal, E., Seydel, E.R., & van de Laar, M.A.F.J. (2009). Participation in online patient support groups endorses patients' empowerment. *Patient Education and Counseling, 74*(1), 61–69.

Van Wormer, K. & Thyer, B.A. (2009). *Evidence Based Practice in the Field of Substance Abuse. A Book of Readings.* Sage, London.

Vanelli, M. (2002). Munchausen's syndrome by proxy web-mediated in a child with factitious hyperglycemia. *The Journal of Pediatrics, 141*(6), 839.

Vartapetiance, A., & Gillam, L. (2012). *Deception detection for the tangled web, 41*(1), 34-47.

Vrij, A., & Mann, S. (2001). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology, 15,* 187–203.

Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest, 11*, 89–121.

Vrij, A. (2015). A cognitive approach to lie detection. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Deception detection: Current challenges and new approaches* (pp. 205–229). Chichester, UK: Wiley

Wallace, P. (2015). *The psychology of the internet* (2nd ed). UK: Cambridge University Press.

Wang, L., & Li, L. (2013). Automatic Text Classification Based on Hidden Markov Model and Support Vector Machine. In Z. Yin, L. Pan, & X. Fang. *Advances in Intelligent Systems and Computing*. In Proceedings of The Eighth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), HuangShan, China (pp. 217-224). London: Springer.

Ware, J. (2011). Cancer networks. In G.A. Barnett (Ed.), *Encyclopedia of Social Networks* (Vol 1) (pp. 109-112). London: SAGE.

Weckerle, A. (2013). *Civility in the digital age: How companies and people can triumph over haters, trolls, bullies, and other jerks.* Indianapolis: Que Publishing.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. New York: The Guilford Press.

Weiss, R. (1974). The provisions of social relationships. In Z. Rubin (Ed.), *Doing unto others* (pp. 48-61). California: Josey-Bass.

Weiss, R. (2004). Treating sex addiction. In: Coombs, R.H. (Ed.), *Handbook of Addictive Disorders: A Practical Guide to Diagnosis and Treatment.* John Wiley & Sons, Hoboken, NJ, pp. 197–232.

White, M., and Dorman, S.M. (2001). Receiving social support online: Implications for health education. *Health Education Research, 16*(6), 693- 707.

Whiteman, N. (2012). *Undoing ethics: Rethinking practice in online research*. Leicester: Springer.

Wiener, M., & Mehrabian, A. (1968). *Language within language: Immediacy, a channel in verbal communication*. New York: Appleton-Century-Crofts.

Witney, C., Hendricks, J., & Cope, V. (2015). Munchausen by Internet and Nursing Practice: An Ethnonetnographic Case Study. *International Journal of Nursing and Clinical Practice, 2*. doi:10.15344/2394-4978/2015/131

Whitty, M., & Joinson, A. (2010). *Truth, lies and trust on the internet.* London, UK: Routledge.

Woodworth, M., Hancock, J. T., & Goorha, S (2005, February). *The Motivational Enhancement Effect: Implications for our Chosen Modes of Communication.* In Proceedings of the 38th Annual Hawaii International Conference on System Sciences, Hawaii, US. Retrieved from: https://www.computer. org/csdl/ proc eedings/hicss/2005/2268/01/22680022a.pdf

Worley, C.B., Feldman, M.D., & Hamilton, J.C. (2009). The case of Factitious Disorder versus Malingering. *Psychiatric Times.* Retrieved from: http://www.psychiatrictimes.com/munchausen-syndrome/case-factitious-disorder-versus-malingering/page/0/1.

Wright, T. (2013, January 10). *Emily Dirr and Warrior Eli* [web log post]. Retreieved from: http://www.warriorelihoax.com/2013/01/10/emily-dirr-and-warrior-eli/

Wright, T. (2012). *A message from Emily* [web log post]. Retrieved from: http://www.warriorelihoax.com/2012/05/15/a-message-from-emily-5-2/

Yamak, Z., Saunier, J., & Vercouter, L. (2016,). *Detection of multiple identity manipulation in collaborative projects*. In Proceedings of the 25th International Conference Companion on World Wide Web, Montreal, Canada, (pp. 955–960). Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

Yi, S., & Baumgartner, H. (2011). Coping with guilt and shame in the impulse buying context. *Journal of Economic Psychology 32*, 458–467.

Yin, D., Xue, Z., Hong, L., Davison, B., Kontostathis, A., Edwards, L. (2009, April). *Detection of Harassment on Web 2.0*. In Proceedings of the 1st Content Analysis in Web 2.0 Workshop, Madrid, Spain.

Zahavi, A. (1977). The cost of honesty (further remarks on the handicap principle). *Journal of theoretical biology, 67*, 603-605.

Zhou, L., Burgoon, J. K., Nunamaker, J., Jay F., & Twitchell, D. P. (2004). Automated linguistics based cues for detecting deception in text-based asynchronous computer-mediated communication: An empirical investigation. *Group Decision and Negotiation, 13*(1), 81-106.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1-59). New York: Academic Press.

Zuckerman, M., & Driver, R. E. (1985). Telling lies: Verbal and nonverbal correlates of deception. In A. W. Siegman & S. Feldstein (Eds.), *Multichannel integrations of nonverbal behavior* (pp. 129–147). Hillsdale, NJ: Erlbaum.

Appendices

Appendix A. Examples of coding from study 1: When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder

Data from online forum REDACTION: Data removed for confidentiality reasons	Open coding: Line-by-line	Focused Coding: Category	Focused Coding: Core Category	Theoretical Code
[redacted]	Witnessing emergency services triggers craving	Environmental stimulus related to emergencies	Episodic triggers	Motivation
[redacted]	Staring at ED and longing to be a patient			
[redacted]	Feeling nervous excitement when ambulance passes			
[redacted]	Any sighting of emergency related services triggers FD			
[redacted]	Being around others helps them not lie	Feeling emotionally		

[redacted]

Becoming overwhelmed by emptiness when alone

Having strong urge for attention/love when alone

Feeling empty when alone

Alleviating void from loneliness with attention

[redacted]

Being alone triggering negative reflection on life

Seeking attention when lonely fills them up emotionally

[redacted]

Helping others rids them of thoughts

Getting attention for being the hero rather than victim

[redacted]

Choosing to be rescuer rather than victim

Wanting to be a rescuer

Taking part in disaster as a rescuer

[redacted]

distressed when alone

Occupying the hero role as opposed to sick role

Recovery Strategies

Treatment

[redacted]

Being a paramedic helps diminish thoughts

[redacted]

Helping others negates need for sick role

[redacted]

Helping others becomes the priority

[redacted]

Eradicating thoughts by helping others

[redacted]

[redacted]

Being a paramedic relieved them of desire for sick role

[redacted]

[redacted]

[redacted]

Removing obstacles to allow them confront issues

Taking control of their enviornment

[redacted]

Confronting issues behind maladaptive behaviour

[redacted]

Retreating to somewhere busy

[redacted]

Seeking change of situation

[redacted]

[redacted]

Retreating to a safe place

[redacted]

Finding somewhere that doesn't elicit FD for respite

[redacted]

Returning home when emotionally strong again

[redacted]

Respite enabling them to use positive coping strategies

[redacted]

[redacted]

[redacted]	Becoming debilitated by fending off FD thoughts	Suicide attempts due to FD	Poor Mental Health Symptoms
[redacted]	FD Exhaustion leading to depression and suicide attempt		
[redacted]	Realising they had a disorder in psych ward		
[redacted]	FD caused suicide attempt		
[redacted]	Discovering disorder due to suicide attempt		
[redacted]			
[redacted]	Having not control of themselves	Suicidal ideation due to FD	
[redacted]	Becoming distressed by feelings		
[redacted]	Contemplated suicide		
[redacted]	Distress leading them to appreciate why others use substances		
[redacted]	Hating themselves		

[redacted]	Believing they would be better off dead		
[redacted]	Despite thoughts could not kill themselves		
[redacted]	Fear of others uncovering their deceit dissuades suicide		
[redacted]	Rationally know how to stop but emotionally can't	Can't stop themselves from enacting FD	Addiction like qualities
[redacted]	Lying before they are consciously aware		
[redacted]	Feeling despondent about solution		
[redacted]			
[redacted]			
[redacted]			
[redacted]	Despite best efforts, contacting those they lied to		
[redacted]	Feeling they have no control		
[redacted]	Wanting attention despite knowing its wrong & wanting attention		
[redacted]			
[redacted]	Experiencing relapse		
[redacted]	Feeling ashamed and out of control		
[redacted]	Blaming self for not wanting to stop enough		
[redacted]	When in rational state wants to stop		
[redacted]	Impulse overpowers them		
[redacted]	Believing they are responsible for behaviour		

[redacted]

Controlling FD is difficult

Finding stopping and enacting both troublesome

[redacted]

Successfully stopping enacting FD for a limited time

Relapsing due to an acute urge

Relapsing due to habituation

Despite feeling guilty compelled to enact FD

Feeling positive about pro/con list

Getting attention is a pro

Feeling guilty is a con

Feeling genuine and not guilty when not lying

Relationships formed are more satisfactory as its genuine

In logical state it's easy to envisage benefits of stopping

During turbulent times difficult to maintain perspective

Questioning if it is akin to addiction

Making direct comparison with perception of addiction

Refusing invite out of guilt from previous lies	Guilt	Feeling of guilt and shame resulting from behaviour
Feeling friendship is fake due to lies		
Feeling remorse deep down		
Feeling extremely guilty on occasion		
Questioning if others feel guilt/remorse		
Most difficult aspect of FD is guilt and anger		
Will always have a sense of guilt		
Affecting because they are caring people		
Not feeling guilty would mean no conscience		
Many believing they are bad/evil rather than ill		
Feeling so ashamed they can disclose FD	Shame	
Feeling physically sickened by their FD		
Not wanting to enact FD due to shame felt		
Shame preventing them having genuine relationships		
Feeling ashamed of past and present behaviour		

[redacted]	Escalating behaviour necessitated withdrawal
[redacted]	Feeling ashamed as they are lying
[redacted]	Despite knowing behaviour is wrong, needing attention
[redacted]	Lying was worth it to fulfil their needs

Appendix B. Examples of coding from study 2: When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder

Data from online forum REDACTION: Data removed for confidentiality reasons	Open coding: Line-by-line	Focused Coding: Category	Focused Coding: Core Category	Theoretical Code
[redacted]	Accused of Munchausen by internet	Falsely accused	Witch hunts	Effect of MbI on community
[redacted]	Accusation came from bonded online community			
[redacted]	Accused hoping for support after stressful doctor's visit			
[redacted]	Disliking moderator who was first to reply			
[redacted]	Moderator questioning veracity of story			
[redacted]	Moderator pointing to discrepancies in experience			
[redacted]	Moderated hoping they would die if a faker			
[redacted]	Moderators friends supporting accusations			
[redacted]	Questioning the legitimacy of brain surgery			
[redacted]	Questioning the lack of details about seizures			

[redacted]	Questioning lack of photos from hospital
[redacted]	Accused wondering why a person would post pics
[redacted]	Questioning how they hadn't suffocated on vomit
[redacted]	Making death threat at accused
[redacted]	Accused retorting they did not need to prove it
[redacted]	Accused defence, they were just seeking support
[redacted]	Accused defence, abuse was breaking rules
[redacted]	Owner protecting Mod, rules don't apply
[redacted]	Threatening with banning if no evidence provided
[redacted]	Demanding scan of receipt from doctor appt
[redacted]	Demanding doctors number to establish veracity
[redacted]	Accused refusing to give evidence
[redacted]	
[redacted]	

[redacted]	Threatening to tell Facebook friends	
[redacted]	rumours were being spread about them	
[redacted]	online	
[redacted]	Stress of accusations and threats causing	
[redacted]	seizure	
[redacted]	Closing account due to accusations	
[redacted]	Blocking members from Facebook account	
[redacted]	Accusation was violent and nasty	
[redacted]	experience	
[redacted]	Questioning if reaction was due to previous	
[redacted]	experiences on forum	
[redacted]	Never returning due to negative experience	
[redacted]		
[redacted]		
[redacted]		
[redacted]		
[redacted]		
[redacted]	Omitting atypical events for fear of	Censoring
[redacted]	accusation	
[redacted]	Fearing being disbelieved despite being	
[redacted]	true	

[redacted]

[redacted] Fear of accusation stopped them posting advice

[redacted] Experience of disorder but feared accusation

[redacted] Bulimia may not be believed on top of atypical experiences

[redacted]

[redacted] Litany of atypical experiences appearing suspicious

[redacted]

[redacted] Fearing others would question their veracity

[redacted] Damaging to make uneducated accusations Risk of false accusations

[redacted] Need understand the syndrome and signs

[redacted]

[redacted] Proposing education about the syndrome

[redacted]

Beneficial to share knowledge of syndrome

Ignorant accusers are hazardous

[redacted]

Risk of casting out genuine members

[redacted]

Not posting about difficult issues for fear of accusations

[redacted]

Other people also not sharing over fears

[redacted]

Fear of being accused discouraging support seeking

[redacted]

People do have incredulous experiences

[redacted]

Not reporting as perceived veracity of story in unreliable

[redacted]

Giving the benefit of the doubt, regardless others may benefit from advice

Resilience

[redacted] Responding as if everyone is genuine

[redacted] Those experiencing the issue may benefit

[redacted]

[redacted] Stories that appear incredulous could be

[redacted] true

[redacted]

[redacted] Accusations offensive to those

[redacted] experiencing issue

[redacted] If suspicious may not respond

[redacted] Responding as if genuine in the event they

[redacted] are need help

[redacted]

[redacted] Someone genuine may benefit from advice

[redacted]

[redacted] Not becoming overly suspicious

[redacted] Avoiding a culture of witch hunts

[redacted] Preferring to offer help despite the risk

[redacted]

[redacted] Preferring to help than risk ignoring a

[redacted] person in need

[redacted]	Suspicious of posts that don't make sense	Inconsistencies/contradictions	Clues which raise suspicion	Detecting MbI
[redacted]	Ultrasound dates being incorrect			
[redacted]	Members appear to be too young			
[redacted]				
[redacted]	Inconsistencies in stories told to multiple members			
[redacted]				
[redacted]	Contradictory dates only clue raising suspicions			
[redacted]				
[redacted]				
[redacted]	Pregnant day after starting fertility drugs			
[redacted]				
[redacted]				
[redacted]	Questioning veracity of such a quick doctor's appointment			
[redacted]				
[redacted]	Due date inconsistent with date of conception			
[redacted]				

[redacted]

[redacted] Questioning veracity of taking fertility drug after conceiving

[redacted]

[redacted] Posts feeling ‘fishy’

[redacted]

[redacted]’

[redacted] Creating stories with numerous issues Dramatic/Atypical

[redacted] Constantly experiencing issues

[redacted] Enjoying attention/advice received

[redacted] Some experience frequent positive events

[redacted] Some experience frequent negative events

[redacted]

[redacted] Series of negative events with no positives

[redacted]

[redacted] Claiming to experience a rare disease

[redacted] Belonging to tiny cohort susceptible to disease

[redacted] Experiencing an incurable disease

[redacted] Constantly grumbling and complaining

[redacted]

[redacted] When others lose interest, change the story

[redacted] Incurable disease goes into remission

[redacted]

[redacted]

[redacted] Investigating suspected member | Authoritative clues | Clues which

[redacted] No death notices in paper | confirm

[redacted] No record of member in claimed workplace | suspicions

[redacted]

[redacted]

[redacted] Passing stolen pictures off as their own

[redacted]

[redacted]

[redacted] Investigating suspected member

[redacted] Stolen picture from ultrasound website

[redacted] Picture date predated claimed ultrasound

[redacted]

[redacted]

Appendix C. Examples of memos from study 1: When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder

Sample Memos
Isolation Feeling isolated runs throughout a lot of the member's experiences of FD. However, the context of the isolation runs through a few categories. Firstly, isolation can be a result of FD, in that they feel disconnected from other people because they are lying to them and are not forming relationships that are genuine and based on the truth. For example, some questioned if people would be friends with them if they didn't have these feigned issues. Secondly, feeling isolated can be the cause of enacting FD. The attention received from feigning issues alleviates their isolation through the attention they receive. Thirdly, some members used isolation as a prevention strategy. By withdrawing themselves from society or a situation where they feel they are about to enact FD, they can stop the behaviour. Fourthly, some members described feeling alone in having the disorder, thinking they were the only ones. Meeting others who were experiencing the same issues on the forum gave them a sense of relief and camaraderie. The category isolation needs to be divided up and merged with the larger core categories to reflect this diversity. *The recovery dilemma* The members face a series of dilemmas when it comes to recovery. Only a small proportion seek professional help, despite numerous members wanting to recover and recognising the need for professional help. The reluctance to seek help appears to be mainly out of fear of those they have duped finding out about the lies. It would also mean stopping enacting FD, something that they are afraid of losing because it has been such a safety blanket/coping mechanism for them. This conflict between wanting to recover but also losing their primary coping mechanism is a very strong thread throughout the members experiences. A lot of the members experience emotional torment over their behaviour but still have such an overwhelming need to enact FD that it overrides this torment. Some expressed also being perplexed by this dilemma, they can logically see that the behaviour is harming them and those around them, but they just can't escape the emotional draw to enacting FD. This contradicts traditional notions that those with FD don't want to recover. From the outside, it

does look this way, but by delving deeper the dilemma that underlies this apparent 'blaséness' tells a different story. It's important that this a strong theoretical category in the paper because it departs from traditional theories. Also, there is the hope that is could help change attitudes towards people with FD as being completely nonchalant about their deceptive behaviour that looks completely selfish.

Motivation category cont.

Motivation category is still too broad and disparate. There are about 30 sub-categories within it. This is because a large variety of possible motivations were proposed by the members so that in some instances a category may only be made up of one response. They need to grouped together so they can be easily understood. Therefore 13 categories have been created to account for the original 30 sub-categories with no less than 4 responses per category. There is one exception, 'no reason', there were 2 people in this group who were perplexed by their behaviour. It was important to keep this category because it could reflect the sentiment of the third of people who did not explicitly state a reason for their behaviour, however we can't know this for sure. But all perspectives on motivation should be given a voice regardless of the numbers.

Motivation cont…cont…

Motivation is proving to be deceptively tricky to organise. With regards to presenting it in the paper, more reorganising needs to be done. Within the 13 categories there appears to be another distinction, between motivation that is ever present and motivation to engage in FD which is sporadic. So, in the latter case there are certain situations/event that will spark a need to enact FD and in the former, there is a continuous motivation associated with FD. These could also be split into 'after the fact' (benefits of enacting FD) e.g. duping delight, validation of existing disorder and 'before the fact' (what pushed them over into enacting FD) e.g. emotional distress, life changing event, but then some sub-categories don't neatly fall e.g childhood issues. So, the original proposal seems to best capture the variety sub-categories and it will also make the motivation category more 'user friendly' in the paper because in its current state there are too many unique categories.

Malingering

Because malingering is not a disorder it's important to make sure instances of people purely feigning for external gains is curtailed. However, this posed an interesting issue. Some of the members were feigning for external benefits, but this seems to be regarded as a secondary benefit rather than the primary goal. Therefore, it is important to include this in the motivation

category because although not the intended goal, it is a ‘perk’ of feigning. In fact, to confuse the instance more, some were adding an extra layer of feigning to obscure their original deceptive behaviour.

Testament of desire to recover

A large variety of recovery strategies have been described in the absence of seeking professional help. These self-recovery strategies are further evidence of the desire to recover. However, it’s difficult to say whether they are just putting a plaster on a wound rather than trying to heal it completely. Some of the most maladaptive ways of trying to recover were isolation, fear of reincarceration and fear of being found out e.g. doctors. The loss of a coping mechanism in the absence of a healthier alternative seemed to be causing even more issues for these people. However, the remaining coping strategies seemed to be healthier with a lot of self-care, self-reflection and getting to know how get their needs, that were previously met through FD, met in a healthier way.

Appendix D. Examples of memos from study 2: When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder

Sample memos
Flamed for raising suspicions Looking back at a handful of categories that haven't found a home, a few members have brought up the issue of being harassed when they suggested that other people were lying. This group seems to be those on the receiving end of other members warning against indiscriminately attacking others suspected of lying without real evidence. Therefore, this stray category can be housed under 'witch hunts' because it is further evidence of members trying to deter witch hunts within the community. *Despondent – the pragmatic approach* Most users have a very pragmatic approach to dealing with MbI, simply put, 'report and ignore'. The despondency arises from the recognition that there is not much moderators can do, despite their best efforts, to prevent cases of MbI appearing in the community. Furthermore, there are no deterrents in place to stop MbI users re-joining the group under a new name and engaging in deception again (see Feldman report on case brought to court as an alternative method). Directly attacking suspected cases of MbI also poses problems as false accusations can occur (instances of people reportedly being bullied because they were suspected, can be fitted into the narrative here as evidence). Therefore, most members advocate reporting a suspected case to the moderator and then moving on. In terms of prevention, they also take a pragmatic approach, don't become overly invested in others online. However, although certainly sensible, it is sad to see that those who are members of online support groups feel the need to remain emotionally distant from one another when support groups should foster the exact opposite environment. *Investment and commitment????* Thinking through where to house this category. The description given to it is: To begin with the members believed that maintaining these fake stories took a significant investment and commitment on the part of the faker, requiring a certain amount of research to pass themselves off as having a certain issue. Some of the members even applauded the dedication it took. This has been placed in 'identifying MbI' but it just doesn't make sense here as

although the stories are elaborately constructed, this in and of itself is not helpful in identifying MbI as the elaboration makes the stories even more convincing. Therefore, it serves the exact opposite function in helping to identify MbI. It seems to best reflect a type of MbI that is very sophisticated rather than being a cue. Thinking that the best option is to house this category with Trolling, obvious cases of MbI. These two categories could be included under the umbrella of 'Types of MbI Behaviour' and represent the juvenile version and sophisticated versions of MbI.

Consequences for the community

'Consequences for the community' is a very large category, spanning a few pages, that needs to be refined further. This category is based an initial cross comparison of the grouped initial codes taken from all forums. These codes were brought together under this category because they all tap into this concern that the members have about MbI ruining their community. On first inspection, it looks like these fears are specifically related to the potential erosion of the community, in fact this could be candidate for renaming the category. Within the category there is constant reference to fears that the presence of fakes will deter others from sharing their stories and seeking help, that it will deter others offering help for fear it's a fake and feeling less trusting of others.

Sinister form of MbI, hard to rationalise motive.

One form of MbI that is particularly sinister is the MbI member befriending someone with the same issues they are feigning. This has appeared under the category of 'exploitation and manipulation'. It seems to be in a league of its own in that it so cold hearted. Members effected by this type describe how the MbI member contacted them to strike up a friendship because of their shared conditions. The progression of the MbI members condition would always be more severe than that of the members, in some instances they feigned death. Given that the befriended member is genuinely experiencing the same issue….it seems a particularly cruel form of MbI given the high emotional investment. It also seems unnecessarily cruel given that they seek these people out and could just as easily received attention/sympathy from the broader forum without becoming that emotionally entangled with someone else who is very vulnerable themselves. It is troubling why anyone would be this cruel when they could have their needs fulfilled in a less evil/harmful way.

Possible explanation for sympathy? Keep in mind

Surprising finding was the sympathy that the members have towards MbI perpetrators given the disruption they were causing to the online community. Furthermore, it is also unexpected

given the amount of stigma towards FD, even from medical professionals. While there may be straight forward explanations for this goodwill e.g. members of online communities can maintain a balanced perspective given that they are not in FtF contact with the perpetrator and it's not 'costing' them too much compared to those who must deal with FD up close and personal. However, one possible alternative explanation is that those expressing sympathy have FD themselves. There is nothing that can be done to assess this risk, other than to say that named MbI perpetrators who had contributed to the conversations that were collected, were removed from the data. It is important to include this in the limitations but also balance it out with a defence.