Head Office: UNIVERSITÀ DEGLI STUDI DI PADOVA

Department of General Psychology

Ph.D. COURSE IN: **PSYCHOLOGICAL SCIENCES**
SERIES: **XXX°**

The assessment of mood disorders: "New methodological perspectives for differential diagnosis"

Coordinator: **Ch.mo Prof. Giovanni Galfano**

Supervisor: **Ch.mo Prof. Giulio Vidotto**

Co-supervisor: **Ch.mo Prof. Gianmarco Altoè**

Ph.D Student: **Francesca Serra**

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

ABSTRACT (Italian version)

Il presente lavoro di tesi si propone di offrire nuove prospettive metodologiche nell'assessment dei disturbi dell'umore, con l'obiettivo principale di suggerire alternative efficaci alla valutazione dell'episodio depressivo maggiore, nell'ottica di sostenere la diagnosi differenziale di diverse forme di depressione.

I disturbi dell'umore sono il più frequente disturbo mentale e la loro incidenza è aumentata negli ultimi decenni, diventando uno dei più significativi problemi socio-sanitari. Perdita del lavoro, divorzio, difficoltà nel crescere i figli e abuso di sostanze sono solo alcuni dei gravi rischi associati ai disturbi dell'umore. Il suicidio è la più tragica delle conseguenze. Il decorso di questi disturbi così come la loro prognosi sono strettamente legati alla corretta diagnosi e al tempestivo trattamento. Purtroppo, attualmente è molto alto il rischio di diagnosi non corretta, con gravi ripercussioni sul trattamento e quindi sul decorso della malattia. In particolare l'episodio depressivo maggiore viene troppo spesso classificato in un solo modo e senza specificazioni, nonostante le possibili diverse configurazioni di sintomi che lo caratterizzano. Come tale esso viene trattato con farmaci antidepressivi, che in alcuni casi (ad esempio la depressione agitata) possono non solo aumentare i sintomi di agitazione, ma anche aumentare il rischio di suicidio.

La fase di assessment riveste un ruolo cruciale in vista di un trattamento adeguato del disturbo. I medici dopo aver raccolto il maggior numero possibile di informazioni sul paziente, devono formulare ipotesi diagnostiche in breve tempo per pianificare interventi clinici efficaci. La qualità della valutazione clinica è fondamentale sia per la diagnosi che per il trattamento.

Il Formal Psychological Assessment (FPA; Spoto, 2011; Spoto, Bottesi, Sanavio & Vidotto, 2013) si configura come una metodologia che unisce i vantaggi delle interviste

semi-strutturate e dei self-report, cercando di superare i loro limiti. Infatti l'approccio metodologico dell'FPA permette la costruzione di strumenti:

- In grado di restituire delle informazioni qualitative, relative ai sintomi del paziente, che vanno oltre lo score numerico.

- In grado di differenziare pazienti che ottengono lo stesso punteggio al test, ma che hanno risposto a item diversi, e che hanno quindi configurazioni diverse di sintomi.

- Adattivi (come le interviste semi-strutturate) che permettono di indagare le aree sintomatologiche del paziente e di approfondirle.

- Di rapida somministrazione come i questionari self-report.

Nel progetto svolto in questi tre anni all'Università di Padova, sono stati utilizzati i concetti dell'FPA in diverse fasi. In una prima fase è stata svolta un'analisi metodologica dei questionari self-report più utilizzati nel campo della depressione per esplorare la loro capacità di indagare tutti i sintomi dell'episodio depressivo maggiore. La ricerca si è basata sulle relazioni tra gli "item" e i "criteri diagnostici" per la depressione, in linea con la metodologia dell'FPA. Nella seconda fase, è stato costruito un nuovo questionario di 41 item sulla base di 23 criteri clinici per l'episodio depressivo maggiore, ricavati dal DSM-5, e dalla diffusa letteratura sulla depressione. Nella terza fase il questionario è stato validato su una popolazione non clinica di 265 individui e su una popolazione clinica di 38 pazienti con episodio depressivo maggiore diagnosticati con depressione maggiore o disturbo bipolare. Il questionario ha mostrato buoni risultati sia per i diversi criteri di validità che per l'affidabilità. Tuttavia, la peculiarità di questo strumento sta nella sua capacità di andare oltre lo score numerico, permettendo di differenziare individui con lo stesso punteggio al test ma che presentano diverse sintomatologie. Questa proprietà è garantita dallo stato clinico del paziente (concetto fondamentale dell'FPA), come principale output del test, ossia dall'insieme di item a

cui l'individuo ha risposto affermativamente con il sotto-insieme di sintomi indagati da quegli item. In questo modo la valutazione clinica non sarà solo legata al livello di depressione ottenuto dallo score, ma dalla configurazione specifica di sintomi manifestati da una precisa persona. Nella quarta fase della ricerca, è stato implementato l'algoritmo computerizzato per il nuovo questionario, in modo da ottenere la forma adattiva dello strumento. Per raggiungere quest'ultimo step, il questionario è stato suddiviso nelle sue tre sotto-scale (affettiva, somatica e cognitiva) corrispondenti ai tre sotto-fattori della struttura fattoriale. Per ogni sotto-scala, attraverso il Basic Local Independent Model (BLIM), modello probabilistico dell'FPA, sono stati stimati i parametri relativi alle probabilità di falso positivo, falso negativo per ogni item e di tutti gli stati clinici della struttura. É stata utilizzata una procedura interattiva per massima verosimiglianza, che ha fornito una stima dei parametri e degli indici di fit. Una volta testato sui dati reali, la forma adattiva dello strumento permette una somministrazione più rapida ed efficiente. Infatti, gli item a cui l'individuo dovrà rispondere dipenderanno dalle risposte precedentemente date, in un processo che imita l'intervista semi-strutturata, evitando possibili inferenze logiche del clinico. Il nuovo strumento per l'assessment della depressione chiamato QuEDS (Quantitative and Qualitative Evaluation of Depressive Symptomatology) rappresenta quindi un supporto per lo psichiatra o lo psicoterapeuta, in quanto offre la possibilità di distinguere i sintomi depressivi di ogni individuo al di là dello score ottenuto al test, e permette di somministrare solo gli item legati alla sua sintomatologia seguendo il flusso logico di domanda-risposta. Dunque due pazienti che ottengono lo stesso punteggio al test, indice dello stesso potenziale livello di depressione, potranno essere trattati comunque in accordo con i loro sintomi; infatti aver risposto allo stesso numero di item non significa aver risposto agli stessi item. In particolare è noto che l'uso di farmaci antidepressivi non è sempre consigliato nella depressione. Esistono infatti le depressioni "miste", così

definite da moltissimi autori, perché caratterizzate sia da sintomi depressivi che da sintomi maniacali (come agitazione, angoscia, irritabilità, insonnia, labilità emotiva). Due esempi di depressione mista sono la depressione agitata e la depressione con fuga delle idee, in cui i farmaci antidepressivi non solo aumentano la componente eccitatoria (quindi i sintomi maniacali) peggiorando il decorso della malattia ma, problema ancora più grave aumentano il rischio di suicidio. Per questo motivo capire tutta la sintomatologia depressiva risulta fondamentale nella pratica clinica. L'ultima parte del progetto di questi tre anni, è stata svolta in Inghilterra, in collaborazione con le Università di Cardiff e Worcester, in particolare con Il Bipolar Disorder Research Network (BDRN). I dati del BDRN utilizzati in questa ricerca comprendono 3750 pazienti con disturbi dell'umore divisi nei tre sotto-gruppi: Disturbo Depressivo Maggiore (MDD), Disturbo Bipolare di tipo I (BD-I) e Disturbo Bipolare di tipo II (BD-II); nel 29,3% dell'intero campione era presente un episodio di depressione agitata, in particolare la depressione agitata era più presente nel disturbo bipolare, soprattutto BD-II. Inoltre i pazienti con depressione agitata avevano più comorbidità con disturbo di panico e con abuso di sostanze, facevano maggior uso di psicofarmaci, e soffrivano di maggiori episodi misti durante l'arco di vita. La depressione agitata era correlata ai tentati suicidi durante l'arco di vita e all'ideazione suicidaria durante l'episodio affettivo. Questi risultati confermano e rafforzano le indicazioni di diversi altri studi svolti su campioni clinici meno ampi. Il riconoscimento e la diagnosi differenziale della depressione mista è essenziale per evitare una diagnosi scorretta e un successivo trattamento pericoloso. La costruzione di strumenti di supporto al medico, che siano in grado di restituire la configurazione di sintomi del paziente e di garantire maggiori informazioni cliniche può diventare un punto di forza nella pratica clinica. Lo strumento presentato in questo lavoro, rappresenta un passo avanti in questa direzione; tuttavia per permettere una diagnosi differenziale dell'episodio depressivo questo primo step ha

bisogno di essere accompagnato dall'esperienza e la consapevolezza del clinico nel campo dei disturbi dell'umore, e soprattutto dallo sviluppo di ulteriori approfondimenti nel contesto metodologico. Infatti come i dati dimostrano, riuscire a catturare i sintomi di una depressione mista risulta un'impresa ardua sia dal punto di vista clinico che dal punto di vista metodologico per quanto concerne la costruzione di strumenti adatti ed esaustivi. La questione fondamentale che resta aperta, riguarda la capacità di riuscire ad indagare quei sintomi di componente maniacale che vengono sottostimati dal paziente stesso (come il flusso rapido dei pensieri, la labilità emotiva ecc.) in fase depressiva.

I primi tre capitoli di questo lavoro formano la cornice teorica, e il punto di partenza per la ricerca. Nel primo capitolo sono infatti descritti nel dettaglio i disturbi dell'umore: la prevalenza, la componente genetica, la classificazione dei vari disturbi (Depressione Maggiore, Distimia, Disturbo Bipolare I, Disturbo Bipolare II, Disturbo Ciclotimico, e disturbo a cicli rapidi); inoltre viene descritta la depressione, e in seguito la depressione mista con particolare attenzione alla diagnosi differenziale. Infine viene brevemente spiegato il trattamento farmacologico e le teorie eziopatogenetiche della depressione. Nel secondo capitolo viene descritto l'assessment, quindi gli strumenti maggiormente utilizzati con i loro punti di forza e di debolezza; vengono inoltre descritti la batteria CBA 2.0 (Cognitive Behavioural Assessment 2.0), e l'assessment adattivo. Il terzo capitolo è dedicato alla spiegazione dell'FPA a partire dalle teorie matematiche sulle quali si fonda fino alla realizzazione del metodo nel contesto clinico. I capitoli 4, 5, 6, 7 descrivono le quattro ricerche principali di questo progetto di dottorato. Infine nel capitolo 8 sarà presentata la discussione finale dell'intero percorso.

ABSTRACT (English version)

This dissertation work aims to provide new methodological perspectives in the assessment of mood disorders, with the main task of suggesting effective solutions for the evaluation of major depressive episode (MDE), in order to support the differential diagnosis of different forms of depression.

Mood disorders are the most prevalent of all mental health diagnoses and their incidence has increased in recent decades, becoming one of the most significant public health problem. Many people fail to go to school or university, lose their jobs, lose their partner and friends, and may commit suicide. The course of these disorders as well as their prognosis are closely related to proper diagnosis and well-timed treatment. Despite this, the risk of misdiagnosis is currently high, with serious consequences both for the current episode and for the course of the illness. In particular, the MDE is often classified without specification, also if there are different possible configurations of symptoms that characterize it. Thus, MDE is almost always treated with antidepressant drugs that in some cases (e.g., agitated depression) may increase the symptoms of agitation and the risk of suicide.

The assessment phase plays a crucial role for the proper treatment of the disorder. Physicians after collecting patient's information, need to formulate diagnostic hypotheses in a short time to plan effective clinical interventions. The quality of clinical evaluation is crucial for both diagnosis and treatment.

Formal Psychological Assessment (FPA; Spoto, 2011; Spoto, Bottesi, Sanavio & Vidotto) is a new methodology able to maximize the benefits of both semi-structured interviews and self-reports, trying to overcome their limitations. Indeed, the FPA methodological approach allows the construction of:

• Tools that are able to provide qualitative information about patient symptoms that goes beyond the numerical score.

• Tools that are able to differentiate patients who obtain the same score, but replied to different items, and therefore have different symptoms configurations.

• Adaptive tools (as semi-structured interviews) that allow investigating and deepening the patient's symptoms.

• Rapid administration as self-report questionnaires.

In this three-year project at the University of Padua, the FPA concepts have been applied to achieve different aims. In a first phase, a methodological analysis of the most used self-report questionnaires of depression was carried out to explore their ability to investigate all the symptoms of the MDE. The research is based on the relationship between "items" and "diagnostic criteria" for depression, in line with the FPA methodology. In the second phase, a new self-report questionnaire of 41 items was built on the basis of 23 clinical criteria for MDE from DSM-5 and literature. In the third phase, the same questionnaire was validated on non-clinical sample of 265 individuals and clinical sample of 38 patients with MDE diagnosed with major depression or bipolar disorder. The questionnaire provided good results both in terms of validity and reliability. However, the strength of this tool stands in its ability to go beyond the numerical score, allowing to differentiate individuals with the same score but with different symptoms and possibly different severity of the episode. This property is assured by the patient's clinical state (the fundamental concept of FPA) as the main output of the test, which is the set of items the individual replied affirmatively with the subset of symptoms investigated by those items. In this way, clinical evaluation will not only be related to the level of depression obtained from the score but also to the specific configuration of symptoms manifested by the individual. In the fourth phase of the research, the computerized algorithm was implemented in the new questionnaire to

obtain the adaptive form of the tool. The questionnaire was subdivided into its three sub-scales (affective, somatic and cognitive) corresponding to the three sub factors of the factorial structure. For each sub-scale, through the probability model of the FPA (i.e. the Basic Local Independent Model; BLIM) the false negative, false positives for each item and all the clinical states of the structure were estimated. An interactive procedure was used with maximum likelihood, which provided an estimate of parameters and fit indexes. After being tested on real data, the adaptive form of the tool allows faster and more efficient administration. Indeed, the items to which the individual will respond will depend on previous responses, in a process that mimics the semi-structured interview, avoiding possible logical inferences of the clinician. The new tool called quantitative and qualitative evaluation of Depressive Symptomatology (QuEDS) can be a support for clinicians; in fact, it differentiates the individual's depressive symptoms beyond the score and allows administering only the items related to its symptomatology following the logical flow of question-answer. Thus, two patients who obtain the same score on the test can be treated differently according to their symptoms, since answering the same number of items does not mean responding to the same items. In particular, it is well known that the use of antidepressant drugs is not always recommended in depression. There are "mixed" depressions, as defined by many authors, because they are characterized by both depressive symptoms and manic symptoms (such as agitation, anguish, irritability, insomnia, mood lability). Two examples of mixed depression are agitated depression and depression with flight of ideas, in which antidepressant drugs not only increase the excitatory component (manic symptoms) worsening the course of the affective episode but, more seriously, increase the risk of suicide. For this reason, understanding all depressive symptoms is crucial in clinical practice. The last part of this project was carried out in England, in collaboration with the University of Cardiff and Worcester, in particular with the Bipolar Disorder Research Network (BDRN). The

BDRN data used in this research include 3750 mood disorders' patients divided into three subgroups: Major Depressive Disorder (MDD), Bipolar Disorder Type I (BD-I) and Bipolar Disorder Type II (BD-II). The 29.3% of the whole sample had suffered from an episode of agitated depression (AD), particularly AD was more related to bipolar disorder, especially BD-II. Moreover, patients with agitated depression had higher comorbidities with panic disorder and substance abuse, made greater use of psychiatric drugs, and suffer of more mixed states in lifetime. Agitated depression was related to lifetime suicide attempts and suicidal ideation during the affective episode. These results confirm and strengthen the indications of several other studies on smaller clinical samples. The recognition and differential diagnosis of mixed depression is essential to avoid improper treatment with dangerous consequences.

The construction of tools to support clinicians' task providing the patient symptom configuration with more clinical information can become a strength in clinical practice. The tool presented in this work represents a step in this direction; however, to allow differential diagnosis of each MDE, this step needs to be combined to the experience and awareness of the clinician in the field of mood disorders, and especially to the development of further insights in the methodological context. Indeed, as the data demonstrate, the recognition of mixed depression is a difficult task both from a clinical and from a methodological point of view in relation to the construction of suitable and exhaustive instruments. The fundamental issue that remains unclear concerns the ability to investigate the symptoms of manic component that are underestimated by the patient itself (such as racing crowed thought, mood lability etc.) during depression phase.

The first three Chapters form the theoretical framework, and the starting point for the researches. In the first Chapter, mood disorders are described: prevalence, genetic component, diagnostic classification (major depression, dysthymia, bipolar disorder I, bipolar disorder II, cyclotymic disorder, and rapid-cycle disorder); Also depression is

described, followed by mixed depression and the differential diagnosis. Finally, the pharmacological treatment and the etiopathogenetic theories of depression are briefly explained. The second Chapter describes the assessment, therefore the tools most used by the clinician with their strengths and weaknesses; The CBA 2.0 (Cognitive Behavioral Assessment) and Adaptive Assessment are also described. The third Chapter is devoted to the explanation of FPA starting from the mathematical theories to the implementation of the method in the clinical context. Chapters 4, 5, 6, 7 describe the four main researches carried out in this PhD project. To conclude, Chapter 8 will present the final discussion.

CHAPTER 1
Mood Disorders

1. *Global estimates of mood disorders prevalence*

According to World Health Organization 2017 (World Health Organization, 2017; WHO), the proportion of the global population with depression in 2015 is estimated to be 4.4%. Prevalence varies by WHO Region, from a low of 2.6% among males in the Western Pacific Region to 5.9% among females in the African Region. A systematic analysis for the Global Burden of Disease Study (Forouzanfar et al., 2015) showed that the total estimated number of people living with depression increased by 18.4% between 2005 and 2015. This phenomenon could mirror the proportionate increase of depression along the time. The total number of people living with depression in the world is 322 million (Vos et al., 2016), and prevalence rates vary by age, peaking in older adulthood (above 7.5% among females aged 55-74 years, and above 5.5% among males). Depression also occurs in children and adolescents below the age of 15 years, but at a lower level than older age groups. However, this may be due to the fact that depression in children and adolescents manifests with different symptoms (Abela, 2008; Connolly et al., 2017).

More in general mood disorder is among the most prevalent of all mental health diagnoses (Waraich, Goldner, Somers, & Hsu, 2004). It is the most significant public health problem (Bland, 1997) and according to different studies, the high prevalence is significantly associated with its disability, the long duration of illness, high probability of recurrence, the difficulties in diagnosis, the delay in finding the right treatment (Bijl, Ravelli, Van Zessen, 1998; Murray & Lopez, 1996) and, in most cases, with the inefficacy of available pharmacological treatments (see Baldessarini et al., 2013).

Waraich and colleagues in a review of 2004 reported a lifetime prevalence of mood disorders that ranged between 5.5% in Korea and 31.5% in Montreal, while a study of Weissman and colleagues (1996) showed a range of prevalence of 1.5% in Taiwan to 19% in Beirut. The lifetime prevalence was systematically higher for women than for men (Waraich et al., 2004). In particular, in every country rate of major depression were higher among women, but in Bipolar Disorder the female to male ratio was approximately equal at all sites (Cross-National Collaborative Group, 1992; Weissman et al., 1996). For Major Depressive Disorder (MDD) the best estimate rate for lifetime prevalence was 6.7% (Waraich et al., 2004), while the estimate rate for Bipolar Disorder (BD) lifetime prevalence changes depending on the study from 1% to 5% (Bauer & Pfennig, 2005; Judd & Akiskal, 2003; Waraich et al., 2004; Wittchen, 2000; Stefansson, Lindal, Bjornsson, & Guoomundsdottir, 2011). However, it is important to emphasize that the concept of bipolar spectrum includes a significant percentage of sub-threshold cases that are either not diagnosed or diagnosed as major depressive disorder (Bauer & Pfennig, 2005; Frank & Thase, 1999; Merikangas et al., 2011). For this reason, BD with a range of bipolar conditions with less-obvious manifestations is still under-estimated. In other words, investigations of depression and mixed symptoms severity associated with sub-threshold bipolar conditions suggest that this category encompasses clinically significant symptoms that are comparable to people needing treatment for bipolar in outpatient settings, but the manic phases (Especially hypomania of BD-II) are often harder to recognize both by the clinician and by the individual (Goodwin & Jamison, 2007). Furthermore, mood disorders also include the Dysthymic Disorder characterized by chronic low-grade depressive symptoms (McCullough & Clark, 2017; MacQueen et al., 2017). The best estimate rate for lifetime prevalence of Dysthymia was 3.6% (Waraich et al., 2004). Despite the reliability of these results, having a real estimate of the prevalence of mood disorders is a challenge for several

reasons. First because the various conditions of mood disorders can be transient, or recurrent, and in these phases the disorder may not be recognized (Waraich et al., 2004). Second, because the uncertainties of the diagnostic constructs for mood disorders (i.e. the sub-threshold cases of Bipolar spectrum and the Dysthymic disorder). Third, because many people with Major Depressive Disorder develop Bipolar Disorder with antidepressants drugs or, they are reclassified in bipolar spectrum (Goodwin & Jamison, 2007; Holma, Melartin, Holma, & Isometsä, 2008). Fourth, because the differences in the prevalence rates in different studies may be attributed to differences in the method of assessment and in the descriptions used to define depression or more in general mood disorders (Merikangas et al., 2011). Mood disorders are constructs (that cannot be directly measured or observed) developed through inference, hypothesis, deduction, and conjecture.

Suicide is the most tragic consequence of mood disorders (Bauer & Pfennig, 2005; Hasin, Goodwin, Stinson, & Grant, 2005), and up to 20% of Bipolar patients die of suicide (Oquendo et al., 2000). In particular, the odds ratio for suicidal behavior was 6.2 in the people with Bipolar disorder versus control group, while the odds ratio for suicide attempt in people with MDD was 3.1 versus control group (Chen & Dilsaver, 1996) according with US National Epidemiological Catchment Area (ECA). These results mean that about 1 in every 4 or 5 persons with bipolar disorder had made suicide attempts (Batterham et al., 2015; Merikangas et al., 2011). Some data suggest that several factors are associated with an increased risk of suicide: genetic and sociodemographic variables, loss of social and medical support, comorbidity with other disorders or substance abuse, recent environmental adversities (Leverich et al., 2003), and incorrect identification of the illness (e.g. misdiagnosis; Hjorthøj, Madsen, Agerbo, & Nordentoft, 2014).

Mood disorders are associated with high disability, which often compromise critical period of educational, occupational, and social life (Bijl & Ravelli, 2000; Hurley, 2006; Phillips et al, 2009). For all these reasons, an early and adequate treatment could be essential for effective prevention of both suicidal behavior and other complications (Goodwin & Jamison, 2007).

To conclude the incidence of mood disorders has increased in the last decades, and especially in this latter period after the profound socio-economic crisis that has hit the whole world (Editorial Lancet, 2012; Lee, Guo, Tsang, Mak, Wu, & King , 2010; Kupfer, Frank, & Phillips, 2012). Many authors recently reported alarming data on the increase in the incidence of depression and suicides in advanced and developing countries in the world (Pitman, Krysinska, Osborn, & King, 2012; Brent, & Mann, 2005; Rihmer, Kapitany, Gonda, & Dome, 2013; Reeves, Stuckler, McKee, Gunnell, Chang, & Basu, 2012). Some authors argue that because of the severe economic crisis, mood disorders are the most frequent illness in the world, though often unrecognized and sometimes inadequately treated (Editorial Lancet, 2012; Lee et al., 2010; Kupfer et al., 2012; Lozano, Naghavi, & Foreman, 2012). In the last period, suicide is associated with a mood disorder in 90% of cases and with a standard mortality ratio compared to the general population of 20:1 (Baldessarini, Pompili, & Tondo, 2006). Suicide is also the third cause of death in the population aged 15 to 35 (Gunnell & Middleton, 2003).

2. Genetic Epidemiology in mood disorders

In bipolar disorder the risk of became ill for a first-degree relative is around 10 times compared to the risk to became ill in a random person. The risk of became ill in identical twins is about 63 percent (Goodwin, & Jamison, 2007). For major depressive disorder, the risk for first-degree relative to became ill (with major depressive disorder)

is about three times higher than the overall population risk (Goodwin, & Jamison, 2007). The risk in identical twins of person with major depressive disorder is about 34 percent, and the calculated heritability is also about 34 percent. When there are recurrent forms of depression the percentage is higher. Thus considering the data there is a strong genetic component to susceptibility to bipolar disorder and a less strong, though still significant, genetic component to susceptibility to major depressive disorder, especially in the more recurrent forms. Moreover, suicidal behavior is highly familial, and on the basis of twins' studies, is genetic as well. A family history of suicidal behavior is associated with suicidal behavior in the patients especially with mood disorders (Faraone, Tsuang, & Tsuang, 1995). In particular, traits of impulsiveness and aggression in patients and family members are associated with family suicidal behavior, and may contribute to familial transmission of suicidal behavior (Brent, & Mann, 2005).

3. Diagnostic Classification

The classification of mood disorders in DSM-5 (2013) has changed from DSM-IV. The main demarcation in mood disorders is between bipolar and depressive (unipolar) disorders (Regier, Kuhl, & Kupfer, 2013). Several changes may significantly influence how the diagnosis is used in both clinical and research settings. Thus, bipolar disorders range from the classic manic and depressive episodes of psychotic intensity (bipolar I disorder) through recurrent major depressive episodes, alternating with hypomanic episodes (bipolar II disorder), and cyclothymic mood swings. Likewise, depressive disorders include those with psychotic severity, melancholia, atypical features, and dysthymic variants recurring. In general, all mood disorders are characterized by

pervasive dysregulation of mood and psychomotor activity and by related biorhythmic and cognitive disturbances (Akiskal, & Pinto, 1999).

3.1. Major Depressive Disorder

MDD is now located in the "Depressive disorders" section, among the new "disruptive mood dysregulation disorder", persistent depressive disorder, and premenstrual dysphoric disorder. In each chapter, "other specified" and "unspecified" disorder categories allow for diagnosis of individuals who fall short of diagnostic criteria for the core specific disorders (DSM-5, 2013). Most of the criteria for MDD are identical in DSM-IV and DSM-5. The disorder is defined by one or more major depressive episode (MDE) and the lifetime absence of mania and hypomania. To meet criteria for an MDE, it is required that five of nine symptoms are present during the same 2-week period. One of these symptoms must be depressed mood or anhedonia (loss of interest or pleasure). MDD is only diagnosed if an MDE is not better explained by other disorders (schizophrenia, schizoaffective disorder, schizophreniform disorder, delusional disorder, or other psychotic disorder) and if there is no history of hypomania or mania (DSM-5, 2013).

In DSM-5, the MDD diagnosis can be divided into 14 subcategories using severity specifiers. In addition, the MDD section of DSM-5 concludes with a list of specifiers that can be added to diagnoses in this section, including "with anxious distress", "with mixed features", "with melancholic features", "with atypical features", "with mood-congruent psychotic features", "with mood-incongruent psychotic features", "with catatonia", with "peripartum onset", and "with seasonal pattern" (Uher, Payne, Pavlova, & Perlis, 2014).

3.2. Dysthymic Disorder

Dysthymic disorder is distinguished from Major Depressive Disorder by the fact that it is not a sequel to well-defined major depressive episodes, and it is defined by chronic symptoms that are long standing (lasting two years or longer), but do not correspond to all MDD criteria (DSM-5, 2013; Devanand, 2014). Patients often complain that they have always been depressed. People with dysthymic disorder differ from people with MDD because in that depressive symptoms tend to outnumber objective signs of depression (Pakriev, Vasar, Aluoja, Saarma, Shlik, 1998). For this reason marked disturbances in appetite, sleep and libido are uncharacteristic, and psychomotor agitation or retardation is not observed. This all translates into a depression with attenuated symptomatology. Most cases of Dysthymia are of early onset, beginning in childhood or adolescence and certainly by the time patients reach their 20 years (Akiskal, & Cassano, 1998; Pakriev, 1998; Regier et al., 2013).

3.3. Bipolar Disorder

In DSM-5, bipolar disorders are no longer included in a single category with depressive disorders. This new edition of the manual acknowledges for the first time that patients without previous history of bipolar disorder under antidepressant treatment who develop a manic episode of sufficient intensity and duration can be considered as patients with bipolar disorder (DSM-5, 2013). Bipolar Disorders are characterized by severe alterations in mood, emotions, and behaviors, all of which have a variable time span. These mood swings are characterized by the alternation of Maniac/Hypomanic Episodes and Depressive Episodes, for this reason the pathology is defined as Bipolar. Both Mania and Depression have a significant impact on the life of the individual, and are deeply debilitating on the working, social, affective, and family levels. In the DSM-5

the mixed episodes of DSM-IV (the presence of a concurrent manic and major depressive episode) have been removed and replaced with a "mixed features" (De Dios, Goikolea, Colom, Moreno, & Vieta, 2014). The mixed features refer to when depressive or manic/ hypomanic symptoms are present at the same time of mood episodes (depressive or manic/hypomanic; DSM-5, 2013). Bipolar Disorder includes the sub-categories of Bipolar Disorder I, Bipolar Disorder II, Rapid-Cycling Bipolar Disorder, and Cyclothymic Disorder.

> **Bipolar Disorder I:** It is characterized by one or more manic episodes, usually accompanied by major depressive episodes (or manic/hypomanic episodes). Diagnosis of Bipolar Disorder I can exist with the exclusion of schizoaffective disorder, Schizophrenia disorder, delusional disorder, or Disorder of schizophrenia spectrum and other Psychotic disorders with other specification or without specification (DSM-5, 2013). Bipolar I Disorder typically start in the teenage years, the 20s, or the 30s; the first episode could be manic, depressive, or mixed (Judd et al., 2002). There are different way of onset: the mild retarded depression for a few weeks or months, which switches into a manic episode; else a severely psychotic manic episode with schizophreniform features; alternatively, a classic manic episode occurs and it is easier to diagnose. It is also frequent that several depressive episodes occur before the first manic episode (In this case, the diagnosis may be late). Although the overall sex ratio is about one to one, men have on average of more manic episodes and women experience more mixed and depressive episodes (Goodwin & Jamison, 2007). Bipolar I disorder in children is not as rare as previously thought. Childhood affective episode onset is based on irritability features, labile moods, and explosive anger and resulting familial affective loading (Judd et al., 2002).

> **Bipolar Disorder II (and the Soft Bipolar Spectrum):** Different studies conducted showed that between the extremes of classic manic-depressive illness defined by at least one acute manic episode (bipolar I disorder) and strictly defined major depressive disorder (MDD) without any personal or family history of mania, exists an *overlapping group* of intermediary forms characterized by recurrent major depressive episodes and hypomania (Goodwin & Jamison, 2007; Benazzi, 2007). Hypomania refers to a distinct period of at least a few days of mild elevation of mood, positive thinking, and increased energy and activity levels, typically without the impairment characteristic of manic episodes. Bipolar disorder II is characterized by at least one hypomanic episode and at least one Major Depressive Episode (with Maniac Episode Exclusion). In order to diagnose the disorder, it is also necessary to exclude: schizoaffective disorder, schizophrenia disorder, delusional disorder and other psychotic disorders with other specification or no specification (DSM-5, 2013). Because hypomania is experienced as a pleasant, ego-syntonic mood state, persons with bipolar II disorder rarely report it spontaneously (This can be a risk for proper diagnosis). Current data worldwide indicate that bipolar II disorder (with his sub-categories) is actually more prevalent than bipolar I disorder. This certainly appears true in the outpatient setting, where 30% to 50% of persons with MDD have been reported with bipolar II form (Angst, 1998).

> **Cyclothymic Disorder:** it is characterized by frequent short cycles of sub-syndromal depression and hypomania (DSM-5, 2013). People with cyclothymia experience at least 2 years (1 year in children and adolescents) several periods with hypomanic symptoms that do not meet the criteria for a hypomanic episode and numerous periods with depressive symptoms who do not meet the criteria for a major depressive episode (MDE). During this period of 2 years (1 year in

children and in adolescents), hypomanic and depressive periods were present for at least half the time and the individual was not without symptoms for more than 2 months (Perugi, Toni, Travierso, & Akiskal, 2003; Perugi et al., 1998). The course of cyclothymia is continuous or intermittent, with infrequent periods of euthymia. Shifts in mood often lack adequate precipitants (e.g., sudden profound sadness with social withdrawal for a few days switching into happy, social behavior). Circadian factors may account for some of the extremes of mood lability, such as the person goes to sleep in good mood and wakes up early with death thoughts (Perugi et al., 1998).

> **Rapid-Cycling Bipolar Disorder:** it is defined as the occurrence of at least four episodes both depression and hypomania/ mania in a year (DSM-5, 2013). Many such patients require frequent hospitalization because they develop explosive excitement and precipitous descent into severe depression. The disorder is a roller coaster nightmare for the patient, and for family and clinician. Treating these patients is hard and often ineffective. Factors associated with its occurrence include female gender, menopause, temporal lobe dysrhythmias, alcohol, other tranquilizer, stimulant, or caffeine abuse; and long-term, aggressive use of antidepressant medications (Koukopoulus Reginaldi, Tondo, Visioli, & Baldessarini, 2013). Most clinically identified patients are bipolar II women in middle age. Rapid cycling appear less common from bipolar I patients (Calabrese, Rapport, Findling, Shelton, & Kimmel, 2000).

4. Depression: clinical features

The major depressive episode (MDE) is common to both Unipolar Disorder (MDD and Dysthymia) and Bipolar Disorder (in all of their sub-categories). Several features

differentiate different types of MDEs often depending on the type of clinical manifested disorder. "Depressed mood" refers to negative affective arousal, variously described as depressed or anguished, irritable, or anxious. Mood in all of the depressive states is bleak, pessimistic, and despairing (Goodwin & Jamison, 2007). The high perception of pain in many persons with depression is accompanied by an inability to experience normal emotions (Gaillard, Gourion, Llorca, 2013). Patients with MDE may lose the capacity to cry or otherwise have crying spells. The patient often lose the sense of pleasure (anhedonia) he may give up previously enjoyed pastimes. When strong, anhedonia evidences with decreased interest in life and the individual lose all interest in things. In the extreme, patients lose their feelings for their children or spouses, who once were a source of joy (Treadway, & Zald, 2011). Thus, the hedonic deficit in clinical depression might represent a more pervasive inability to experience emotions. The inability of the person to experience normal emotions in depressed patients differs from the schizophrenic patients because the loss of emotions is itself experienced as agonizing; that is, the patient suffers immensely from the inability to experience emotions (Goodwin & Jamison, 2007). Work and social relationships are often severely compromised, and the patient is forced to stop the work and abandon all activities, including leisure activities (Goodwin & Jamison, 2007; Treadway, & Zald, 2011). The essential characteristic of depressive thinking is that the sufferer views everything in an extremely negative light. A deep sense of futility is often accompanied with feeling of guilt. The negative triad of depression includes negative evaluations of the self, the world, and the future (Beck, 1991; Goodwin & Jamison, 2007). Those distorted thinking patterns are clinically expressed as ideas of loss; low self-esteem and self-confidence; self-reproach and pathological guilt; helplessness, hopelessness, and pessimism; and recurrent thoughts of death and suicide. Death ideas are very common: life is not worth living, the patient hopes to die or he plans suicide, makes suicide

attempts, and a high percentage of individuals kill themselves (Goodwin & Jamison, 2007). As reported by Goodwin and Jamison "*The strong tendency to suicide sometimes it accompanies the whole course of the disease, without coming to a serious attempt owing to the incapacity of the patients to decide.... Sometimes the impulse to suicide emerges very suddenly without the patients being able to explain the motives to themselves. Only too often the patients know how to conceal their suicidal intentions behind an apparently cheerful behavior, and then carefully prepare for the execution of their intention at a suitable moment*" (Goodwin & Jamison, 2007 pp. 167). Depression often occurs with somatic symptoms: stomach cramps, vomiting, digestive problems, diarrhea, palpitations, hyperventilation, paresthesia, sweating, hot flashes, tremors, headache, increased heart rate, back pain or muscle pain, decreased energy and fatigue. Changes in appetite and weight are also frequent (Harris, Young, & Hughes, 1984). Sexual dysfunction as well as alterations in circadian rhythms, especially worsening of morning mood are typical features of depression (Goodwin & Jamison, 2007). Depressed people are typically unresponsive to sexual activity or are disinclined to participate in it; this situation could lead to relationship issues. Insomnia is a cardinal sign of depression, often it is characterized by multiple awakenings, especially in the early hours of the morning, rather than by difficulty falling asleep (Tsuno, Besset, & Ritchie, 2005). The attempt to overcome the problem by drinking alcohol or taking sedative-hypnotic may initially work but later aggravates the sleep patterns and insomnia. Although sedatives (including alcohol) effectively reduce the number of awakenings in the short term, they are not effective in the long term because they further diminish stage 3 and stage 4 sleep. They tend to prolong the depression (Benca, & Peterson, 2008). Sleep is perceived by the patient, however, not restorative, because he awakens as if he had not slept for nothing and rested at all (Goodwin & Jamison, 2007). Some patients on the contrary exhibit hypersomnia and have difficulty waking

up in the morning (Billiard, Dolenc, Aldaz, Ondze, & Besset, 1994). In depression, psychomotor disturbances refer of changes in the expression of mental and emotional activity. Psychomotor retardation can be the core of pathology in mood disorders in which the patient experiences inertia, being unable to act physically and mentally. Sometimes depression what patients describe as being "down" can be understood in terms of moderate or extreme psychomotor slowing (Bennabi, Vandel, Papaxanthis, Pozzo, & Haffen, 2013). This may occur with the simple motor slowdown, but it is more often associated with a slowing down of the conception, the speech, and the difficulty of concentration. The same patients who experience psychomotor retardation or other depressed individuals can manifest psychomotor agitation. Agitation clinical picture occurs more frequently with symptoms as restlessness, irritability, crying, anguish, incapacity to sit still, repetitive activity such as pacing up and down, wringing of hands, or even biting nails and/or lips (Benazzi, 2004a, 2004b; Goodwin & Jamison, 2007; McGuffin, Farmer, & Harvey, 1991; Sobin, & Sackeim, 1997). Agitated depression especially in bipolar disorder can assume the features of "mixed state" (Goodwin & Jamison, 2007; Swann et al., 1993).

4.1. Depression with mixed features

Mixed states are mood episodes that have typical symptoms of both depression and mania. It has traditionally been considered a distinct episode from both the depressive and the manic episode (Goodwin & Jamison, 2007; DSM-IV, 2000). According to the last Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, 2013) MDE with mixed features requires the simultaneous presence of at least three of manic/hypomanic symptoms and a full Major Depressive Episode. Many authors had proved it is a very rare feature with the lack of satisfaction of these criteria (Faedda,

Marangoni, & Reginaldi, 2015; Koukopoulos, & Sani, 2014; Maj, Pirozzi, Magliano, & Bartoli, 2003). On the contrary, many studies have shown the presence of episodes of dysphoric mania and agitated depression that do not match the description of major depressive episode with mixed features of the DSM-5 (Dilsaver, Chen, Swann, Shoaib, Tsai-Dilsaver, & Krajewski, 1997; Swann et al, 1993). Particularly according to Sani and colleagues (2014) this change in DSM-5 is able to describe the manic episode with mixed features but is not sufficiently satisfying for the definition of mixed depression (e.g. agitated depression). According to several studies (Benazzi, 2004a; 2004b; Sobin, & Sackeim, 1997; Sani et al., 2014), symptoms of mixed depression that occur most frequently are agitation, irritability, anguish, insomnia, racing or crowded thoughts, mood lability or marked reactivity, talkativeness, dramatic description of suffering or frequent spells of weeping (in addition there are the other common symptoms of the major depressive episode). Mixed features in depression can also manifest with the flight of ideas: thinking processes and perception are accelerated, experienced as unusually sharp; the patient may speak with such pressure that associations are difficult to follow (Goodwin & Jamison, 2007). Kraepelin (1913; 1921) described patients with mixed features as suffers of extreme mental anguish, including the terrifying racing thoughts and feelings. Kraepelin (1913; 1921) classified agitated depression as a result from the combination of opposite polarity of symptoms: mood and thought in depressive polarity and activity in manic polarity. In the same way, Kraepelin classified depression with flight of ideas as a result of negative polarity of mood and activity and positive polarity of thought. In his view, it was enough to have one of the three components (psychomotor activity, mood and thinking) in manic polarity to have mixed state (Akiskal & Benazzi, 2004).

A critical issue exists about Mixed Depression related to its specific characteristics, its prevalence, and its ratio in the different disorders (major depressive disorder, bipolar

disorder I and II). In recent years, this form of depression has drawn the attention of a great number studies (Akiskal, & Benazzi, 2004; Benazzi, Koukopoulos, & Akiskal, 2004; Koukopoulos, Sani, Koukopoulos, Manfredi, Pacchiarotti, Girardi, 2007; Swann et al., 1993), which suggest that it is not at all a rare observation and underscore the necessity of a corrected diagnosis in order to avoid erroneous treatment. Concerning the possible variables associated with AD, some studies correlated the AD with lower age at onset of the mood disorder, more Bipolar disorder, female gender, longer duration of illness, more MDE recurrences, more MDE symptoms (suggesting more severity), more patients with atypical features specifiers (i.e. mood reactivity), and more family history of bipolar disorders. Compared to patients with non-agitated depression, they had a longer time to 50% probability of recovery from the index episode, were more likely to receive standard antipsychotic drugs during that episode, and spent more time in an affective episode during the observation period. Moreover, psychomotor agitation and suicidal ideation were found to be correlated in many studies (Andreasen and Grove, 1982; Kendler, Eaves, Walters, Neale, Heath, & Kessler 1996; Korszun et al., 2004; Maj et al., 2003; Raskin, Schulterbrandt, Reatig, & Mckeon, 1969; Sullivan, Kessler, & Kendler, 2002). Thoughts of death are often linked with mixed depression; the patients wish to die or plan, attempt and, relatively frequently, die by suicide (Goodwin & Jamison, 2007; Olin, Jayewardene, Bunker, & Moreno 2012; Bocquier, Pambrun, Dumesnil, Villani, Verdoux, & Verger, 2013). All these information could play an important role in the assessment and treatment of mixed depression.

5. *The importance of Differential Diagnosis*

Clinical observations and epidemiological studies prove a vast overlap between Bipolar and Major Depressive Disorder. These observations are in line with Kraepelin's view of

mood disorders and his attempt to bring all affective diseases under one rubric (Goodwin & Jamison, 2007). Thus the diagnosis between the various affective subtypes are not fast neither easy. For instance, bipolar disorder can be overlapping on cyclothymic disorder that could persist after the end of manic or depressive episodes. Likewise, evidence indicates that dysthymic disorder may precede major depressive disorder in a third of cases and, crossing from dysthymic disorder to hypomanic/manic episodes has also been described, suggesting that some forms of dysthymia are precursors of bipolar disorder. Moreover, one in four persons with major depressive disorder subsequently develops hypomanic/manic episodes and so should be reclassified as having bipolar disorder. There are many causes for the misdiagnosis of bipolar depression as unipolar depression: for instance patients' lack of insight with regard to manic (especially hypomanic) symptoms (Goodwin & Jamison, 2007); in this regard is well known that depression not only impairs memory but also makes it more likely that memories will focus on past depressions. Thus, even if patients had previously some insight about manic symptoms, they often have difficulty recalling those symptoms clearly and accurately. Furthermore, during depression, mania or hypomania may be remembered simply as a good period and therefore patients do not report it spontaneously in the interview with the clinician (Benazzi Helmi, & Bland, 2002; Goodwin & Jamison, 2007). Moreover clinicians often don't investigate about patient's history from family or other significant and clinicians often focus on euphoric symptoms to diagnose mania, and they don't account on dysphoria or irritability which are also symptoms of manic episode (Benazzi, & Akiskal, 2005; Perugi, Akiskal, Micheli, Toni, Madaro, 2001; Sato, Bottlender, Schroter, Moller, 2003). Rihmer and Kiss (2002) reported that patients with bipolar-II disorder are often misdiagnosed and then included as unipolar patients. Thus, the tendency to misdiagnose bipolar-II disorder

as unipolar depression may contribute to the apparently higher suicide rates in unipolar illness.

Many authors compared the symptoms of unipolar depression with those of bipolar depression, the result may be very helpful for psychotherapeutic and pharmacological treatment. The findings showed more symptoms of anxiety, somatic complaints (Beigel, & Murphy, 1971; Greenhouse, & Geisser, 1959) and psychomotor retardation (Goodwin & Jamison, 2007) in Major depressive Episode (MDE) of Unipolar Patients; while there are more symptoms of tension (Goodwin & Jamison, 2007; Vöhringer, & Perlis, 2016), mood lability (Brockington, Helzer, Hillier, & Francis, 1982; Hantouche, & Akiskal, 2005;), irritability (Benazzi, & Akiskal, 2005; Fava, & Rosenbaum, 1999), late insomnia (Goodwin & Jamison, 2007; Oral, & Vahip, 2004), psychotic features (Coryell, & Tsuang, 1985; Mitchell, 2001; Parker et al., 2000) and comorbid substance abuse (Judd et al., 2003; Marneros, 2004) in MDE of bipolar patients. Furthermore many authors assessed different depressive episodes between bipolar I and bipolar II patients. Bipolar II patients have more number of episodes, more rapid cycling and they spend more time in depression (Benazzi, & Akiskal, 2005; Goodwin & Jamison, 2007; Vieta, Gasto, Otero, Nieto, & Vallejo, 1997); conversely Bipolar I patients have more hospitalizations, irritability, and psychotic features (Goodwin & Jamison, 2007; Serretti & Olgiati, 2005; Vieta et al., 1997). Regarding the prevalence of suicide attempts, different findings have been reported. Some studies show greater suicidal behavior in bipolar II (Goldring, & Fieve, 1984; Rihmer, & Pestality, 1999), other studies suggest suicide attempts are the same in the two disorders (Coryell, Keller, Endicott, Andreasen, Clayton, & Hirschfeld, 1989; Vieta et al., 1997).

Considering the assessment phase, mixed depression have a significant clinical relevance in mood disorders and may occur in both bipolar and unipolar disorder (Akiskal, Benazzi, Perugi, & Rihmer, 2005; Benazzi, Helmi & Bland, 2002).

Nevertheless, there are limits of information of mixed depression and therefore it is already an underestimation of the consequences, which could result in misdiagnosis and inappropriate/wrong treatment, often with very dangerous outcomes both for the course of the illness and for the suffering of patients (Akiskal et al., 2005; Bocquier et al., 2013). In particular, treatment with antidepressant drugs in agitated depression (AD) could worsen the excitatory symptoms resulting in the failure to relieve the patient's pain (Akiskal et al., 2005; Koukopoulos, & Koukopoulos, 1999; Vázquez, Tondo, Undurraga, & Baldessarini, 2013). Indeed, it has been reported that antidepressants monotherapy in AD might increase psychomotor agitation. Moreover, concerns have been reported about the possibility that the antidepressant administration in the agitated depression could increase the risk of suicide (Akiskal et al., 2005; Baldessarini et al., 2006a; Koukopoulos, & Koukopoulos, 1999; Vázquez, Tondo, Undurraga, & Baldessarini, 2013).

These observations suggest that much of the unipolar spectrum might be "soft bipolar". The clinical significance of these considerations are of clinical relevance especially as far as it is concerned the switches in polarity and the resulting clinical and pharmacological treatment of various types of depression (Goodwin & Jamison, 2007). Irritability may be a good marker of depression with mixed features, a view consistent with that of others who have found high rates of irritability and anger attacks associated with these states (Akiskal, & Benazzi, 2003; Koukopoulos et al., 2007; Sani et al., 2014). According to several studies, there is a significant greater proportion of Agitated Depression in Bipolar Disorder than in MDD (Benazzi, 2004a; Benazzi et al., 2004; Koukopoulos et al., 2007; Takeshima, & Oka, 2013).

This awareness can help the clinician to avoid the all-too-common misdiagnosis of agitated depression and other depressive/manic symptoms, a mistake that can lead to the almost always frustrating treatment decision and the administering an antidepressant in

the absence of a mood stabilizer (Baldessarini, Tondo, Davis, Pompili, Goodwin, & Hennen, 2006b; Baldessarini et al., 2013).

Anyway, the data suggest greatly elevated suicide rates in both unipolar and bipolar disorders in comparison with other psychiatric diseases (Sharma, & Markar, 1994; Harris, & Barraclough, 1997).

The most prudent approach is perhaps to give clinicians the opportunity to maximize the assessment phase of each case. The result of a good assessment and therefore of a correct diagnosis is the possibility of treating the individual in an effective way. The diagnosis of a type of depression cannot be accomplished by a checklist: The DSM-5 diagnostic criteria for major depressive episode provide only a general guide. Only after an in depth phenomenological approach can a clinician ascertain diagnosis of a specific affective episode and choose the most appropriate treatment.

6. Etiopathogenetic theories of depression

Hippocrates (460-377 AC) considered depression to as the result of excessive secretion of "black bile" and therefore gave it the name Melancholia, a term that many scientists prefer to name of depression. The Hippocrates intuition could be considered the first biological hypothesis of depression.

The formulation of the first neurobiological hypothesis of depression based on experimental evidences, was the "monoaminergic hypothesis". It still continues to have its validity, although revised by new discoveries (Schildkraut, & Kety, 1967; Serra, & Fratta, 2007). It is based on the studies on the mechanism of the antidepressant effect of the serendipitous discovery of imipramine and MAO-inhibitors and the depressant effect of reserpine (a drug used in the treatment of hypertension, which cause severe depression).

The observation that imipramine potentiates noradrenaline and serotonin transmission by inhibiting the reuptake of these neurotransmitters and that MAO-inhibitors enhances monoamine activity by inhibiting the enzymatic degradation, led to the hypothesis that depression should be associated to a reduced activity of these monoamines (NA and SE) in the SNC. In keeping with this hypothesis, reserpine, which causes severe depression, depletes monoaminergic neurons from their neurotransmitters (NA, SE, DA).

The discovery that antidepressant drugs potentiate also the dopaminergic transmission (See Demontis, Serra, & Serra, 2017) suggested that depression may be associate also to a reduced dopamine activity.

More recently, it has been suggested that depression could be a consequence of a neurodegeneration phenomenon and that the antidepressant and/or the mood stabilizing effect of the drugs should be attribute to their ability to promote neurogenesis (Schloesser et al., 2015).

A detailed description of the various neurobiological theories of depression is not part of this work, thus we will be focused on relevant psychological theories.

Among the "psychological" theories proposed in the last century to explain the psychological mechanisms underlying depression, the most acclaimed cognitive behavioral matrix theories are Beck's "**hopelessness theory**", and Seligman's "**learned helplessness theory**". Both have led the development of today's widely used psychotherapeutic techniques.

7.1. Beck's Theory

Beck's cognitive model has been described in numerous publications (Rush, & Beck, 1978; Beck, 1991; Beck, 2005). At the basis of its descriptive model, there are the persistent structural representations of human experience, called **schemes** that guide the identification, interpretation, categorization and evaluation of experience. These

schemes are structurally rigid, impenetrable and absolute; their content is a distorted representation of the experiences.

The schemes are quickly triggered by a whole series of stimuli and once turned on dominate the information evaluation system. The dominance of the negative assessment also interferes with the assessment of positive events.

Another important aspect concerns childhood experiences; the creation of distorted conceptions of oneself is due to early childhood events, such as loss of parents or abandonment, which sensitize the individual to experiment more seriously future losses in adolescence and adulthood.

These cognitive distortions are triggered by adverse events that, associate with these specific cognitive vulnerabilities, result in systematic cognitive distortions (Clark, & Beck, 2010).

Because of repeated activations, the negative self-scheme, acquires a coherent and elaborate organization that over time turn on easily through a variety of modest, stressful events. Cognitive errors due to these fixed patterns in the individual mind are very important in the development and maintenance of depressive disorder.

In Beck's theory, subjects are considered more vulnerable when are guided by depressive schemes, in which personal value is related to perfectionist standards or with other's approval. When they are faced with negative events, they experience a lack of control. These individuals tend to consider themselves responsible for adverse events and failures (e.g. social relationships, work levels, etc.) due to poor personal value and poor personal skills. A negative self-view is a central feature of subjects who feel depressed; they tend to distort the actual information about their skills in different areas.

At the base of depression, there would be distorted beliefs and dysfunctional expectations that cause affective reactions and symptomatic cognitive manifestations

(Beck, 2005; Clark, & Beck, 2010). Beck has found that cognitive beliefs and errors affect a "cognitive triad" that includes:

• A negative view of self: in terms of personal value ("I'm a loser", "I'm a failure") and in terms of loveliness and anyone's guess ("no one loves me"; "I'm not a person worthy of love")

• A negative view of the world ("The world is a bad and unhappy place"; "Others take advantage of me"; "Life is unfair to me").

• Negative expectations about the future ("It will never change anything"; "I'll always be a failure")

Concerning cognitive distortions, Beck argues that all of us are continually committed in attributing significance to life events and that in depressed patients, the evaluation of such events is often distorted by different dysfunctional cognitive processes.

7.2. Seligman's theory

The origin of Seligman's theory is based on experimental evidence of animal behavior, and in particular by Seligman's best - known experiment on dogs, described by him in 1972.

Seligman's classic experiment for his learned helplessness theory can be summarized as follows: When a dog is instructed to escape from an electric discharge, which is given to his feet, the animal first has a set of behaviors (defecation, urination, etc.) until he accidentally finds way to escape. At times, the animal becomes more and more capable of escaping from the electric discharge as quickly as possible.

Conversely, when a dog gets electric shock in a situation where there is no way to escape, it shows a behavior that is absolutely different. Indeed, this dog soon stops moving and stands still until the end of the shock; in other words, the dog does not

attempt to escape the electrical shock, but rather seems to accept it passively, even when it can avoid them (Seligman, 1978).

From his experiment, Seligman hypothesized that at the base of depression in humans was, in analogy of the animals, a conviction that he could not do anything in the face of the stressful events of life. Thus, he passively accept the consequences of such events (Abramson, Seligman, & Teasdale, 1978; Miller, & Seligman, 1975; Rosellini, & Seligman, 1975; Seligman, Weiss, Weinraub, & Schulman, 1980).

The subject exposed to stressful events, from which he can not escape, learns that his actions have no power to control and modify such events, developing a sensation of learned helplessness at a cognitive level (Abramson et al., 1978; Seligman et al., 1980). This would also show an attitude of passivity towards the environment, as the subject would learn that his behavior is independent of the result. This attitude tends to generalize to new situations because of the expectations of the impossibility of having any control over the future (Alloy, Peterson, Abramson,, & Seligman; 1984; Seligman, 1978).

7. Treatment of depression

Depression treatment includes psychotherapy and pharmacotherapy. The present work focuses in particular on the relationship between depressive symptoms and the appropriate assessment methodology. For this reason, the treatment will only be mentioned but will not be deepened.

The psychotherapist should be able to understand deeply the type of depressive episode. Then, clinician should have solid knowledge of mood disorders and at the same time be flexible to possible mood changes.

Cognitive Behavioral Therapy has proven to be a clear efficacy and has been indicated (Clark, 2011; National Collaborating Centre for Mental Health. UK, 2010) as an

elective therapy in mild and moderate depression. It is also indicated in combination with medications for serious forms of depression.

Drugs used in the treatment of mood disorders include antidepressant, anti-manic/mood stabilizer and antipsychotic drugs (Goodman & Gilman, 2011).

Antidepressants are used in the treatment of depressive episodes and include tricyclic antidepressants (imipramine, amitriptyline, etc.), which increases monoaminergic transmission by inhibiting the neurotransmitter (Serotonin and Noradrenaline) reuptake by the synaptic cleft; monoamine-oxidase inhibitors (tranylcypromine, etc.), which increases monoamine transmission by blocking their enzymatic degradation. These antidepressants have been discovered "serendipitously" in the late 1950s, and are still the most efficacious treatment of depression. More recently have been introduced in the clinical use the selective serotonin (SSRI such as fluoxetine)-noradrenaline (SNAI such as reboxetine)-dopamine (SDAI such as bupropion) reuptake inhibitors that are a class of the more prescribed drugs today.

However, it should be emphasized that administration of antidepressants in agitated depression or other mixed states worsen the symptomatology, possibly because it intensify the manic component of the disorder (see Serra et al, 2014). In particular, the use of antidepressants in patients with mixed symptoms can increase the excitatory symptoms (a switch from depression to mania) and it may result in a higher risk of suicide (Baldessarini et al., 2006a; Koukopoulos, & Koukopoulos, 1999; Vázquez et al., 2013). Indeed, these forms of depression should be treated with anti-manic/mood-stabilizer or antipsychotic drugs.

The first choice of anti-manic and mood stabilizer treatment is lithium: it is used to treat mania and mixed features and to prevent the recurrences of depressive episodes in the Major Depressive Disorder; lithium is used also to prevent (hypo) manic, mixed and depressive episodes in Bipolar disorder.

Some anticonvulsants and atypical antipsychotics are used in patients who are lithium non-responders, but their efficacy are limited and/or questionable.

Antipsychotics include the first generation or neuroleptics drugs, which are in clinical use since the late 1950s (chlorpromazine, haloperidol, etc.) and the more recent introduced "so called" atypical or second-generation antipsychotics (Clozapine, olanzapine, quetiapine, etc.). These drugs are used in the treatment of severe mixed/manic episodes with psychotic symptoms or high agitation (McIntyre et al., 2014).

The psychiatrist should carefully assess the symptoms before prescribing a drug. Several studies have shown that antidepressants are suitable in some types of depression but not in others. In particular, Goodwin and Jamison (2007) report the results of several studies suggesting the use of a mood stabilizer and (sometimes) an atypical antipsychotic in agitated depression in which antidepressant drugs may worsen the course of the affective episode as stated before.

8. Conclusion

Mood disorder is among the most prevalent of all mental health diagnoses. It is associated with high disability, the long duration of illness, high probability of recurrence, the difficulties in diagnosis, the delay in finding the right treatment, and, in most cases, with the inefficacy of available pharmacological treatments. Consequently, it has a high risk of suicide with a standard mortality ratio compared to the general population of 20:1. Some data suggest that several factors are associated with an increased risk of suicide over the incorrect identification of the illness: genetic and sociodemographic variables, loss of social and medical support, comorbidity with other disorders or substance abuse, recent environmental adversities.

Clinical observations and epidemiological studies prove a vast overlap between Bipolar and Unipolar disorder; in particular, patients with bipolar-II disorder are often misdiagnosed and included as unipolar patients. Major depressive episode (MDE) is common to both Unipolar Disorder and Bipolar Disorder (in all of their sub-categories). The proper treatment of mixed depression is essential for the subsequent proper treatment. Indeed, many studies have demonstrated the ineffectiveness of antidepressant treatments in this form of depression, which may, on the contrary, worsen the symptoms and increase the risk of suicide.

In the light of these observations, the result of a good assessment is essential to treating the individual in an effective way.

In the next section, we will present the assessment phase, describing in detail strengths, weaknesses and progress of the research in the field of psycho-diagnostic evaluation.

CAPTHER 2
Assessment

1. Assessment description

The central role of clinicians conducting assessments should be to answer specific questions and to support the patient in making relevant decisions (Grossberg, 1964). The Assessment, or psycho-diagnostic examination, can be defined as a complex process of collecting, analyzing and processing information. The assessment phase is the first action that clinician has to face when he starts to help a patient and, clinicians must integrate a wide range of data and bring into focus different areas of knowledge.

The assessment is based on a first broad-spectrum evaluation to establish psychiatric treatment and psychotherapy, or to relocate the patient to appropriate interventions to the specific case (Groth-Marnat, 2009).

Clinical evaluation and treatment may have different aims; in fact, psychologists and psychiatrists face many forms of suffering and discomfort. The case formulation is obtained by collecting all necessary data, which allows clinicians to reconstruct the mechanisms and processes underlying the presented disorders, agree on treatment goals, and to identify the most appropriate therapy in an effective way (Serra, Spoto, Ghisi, & Vidotto, 2017).

The assessment consists of a systematic sequence of interconnected phases. The first phase involves collecting data about the patient. It begins with the patient's previous history and records followed by the development of tentative hypotheses and the investigation in more detail. At this point, the clinician conducts an interview and administers a variety of psychological tests. From this information, the clinician starts to make inferences. This step focuses on the development of a wide variety of inferences

about the case, which guides future investigation to obtain additional information that are used to confirm, modify, or negate later hypotheses.

Often, in investigating the validity of an inference, a clinician alters either the meaning or the emphasis of an inference or develops entirely new ones (Groth-Marnat, 2009). The validity of that inference is progressively strengthened as the clinician evaluates the degree of consistency and the strength of data that support a particular result. The central aim of the following step is to develop and begin to elaborate on statements relating to the specific case with a further investigation of the personality traits of the person to better understand and integrate the patient's background. It may include describing and discussing general factors, such as cognitive functioning, affect and mood, and interpersonal-intrapersonal level of functioning. In addition, the clinician analyses the social context. Finally, the crucial phase involves the decision-making and requires that the clinician take into account the interaction between personal and situational variables (Bokhari, & Hubert, 2015). Yet the goal of clinician is not merely to describe the person but rather to develop relevant answers to specific questions.

DSM-5 diagnosis needs to be considered within the context of case specific considerations. An example is how a disorder such as depression may be manifested as a stressful event (i.e., bereavement) or within a bipolar disorder (i.e. as a depressive episode). In fact, a reactive depression should be considered in a totally different way from depression with a genetic basis, and differential diagnosis is also essential for treatment. Likewise, the way to manifest depressive symptoms is often related to personality or culture. In addition to noting the cultural identity of the patient, it is also crucial to carefully consider his personality and social context. Anyway, the individual history would need to be decoded in order to identify the underlying depression.

The information obtained through the Assessment allows creating a hypothesis about the person's clinical features, useful for treatment. When errors in diagnosis do occur,

they have the potential to result in dangerous decisions and wrong treatment (Groth-Marnat, 2009).

For all these reasons, it is crucial to take into account both the nomothetic approach and the ideographic approach. The first one focuses its attention on common aspects of various personalities and therefore tends to categorize individuals according to the psychological disorder in question; the second one analyses the single case, and considers each person unique and different and therefore not classifiable (Diener, & Fujita, 1995). These two approaches are not necessarily mutually exclusive, but can be integrated. In fact, in the case formulation, the evaluation uses patterns, laws and explanations applicable to all people to prepare a scientifically founded and nomothetic formulation; but this is then translated into an ideographic explanation, valid only for that individual, with all its specifics and peculiarities (Hayes, Nelson, & Jarrett, 1987).

To avoid missing crucial information, clinicians should use comprehensive approaches by collecting data from the three main levels of information: the subjective level, the behavioural level, and the physiological level (Sanavio, 2007; Spoto, 2011).

✓ The subjective level covers the information the patient provides during the clinical interview, the structured interview, the diaries, the tests, etc. These information are related to the context in which they are collected, to the truthfulness of the information given and, to the relationship with the psychologist and also to his possible mistakes; when the patient is hostile and uninclined to collaboration, is common making the analysis useless.

✓ The behavioural level (i.e. non-verbal channel) is given by the information that comes from direct observation of the individual's behaviour (e.g. non-verbal behaviour during a clinical interview: facial gestures, posture, tone of voice, etc.). None of this information are neutral as the same observer influences it. More the clinician experience will be, the greater the ability to capture this kind

of information will be. The non-verbal sphere is of fundamental importance in the diagnosis of mental disorders. Negative emotions and social behaviours are important indicators and predictors of the severity of mental illness (e.g. depression) and of its clinical remission that however escape the patient's awareness (Annen, Roser, & Brune, 2012; Fiquer, Boggio, & Gorestein, 2013).

✓ The third level of information is the recording of the psychophysiological activation of the individual (e.g. skin conductance, cardiac frequency, temperature etc.). Although this information is fairly reliable, it is not neutral but depend on the conditions in which the person is at that particular time. These "new" techniques are particularly used in specific disorders such as headache, hypertension (Nicassio, Meyerowitz, & Kerns, 2004), anxiety disorders and phobias (Barlow, 2002).

Measures relative to one or the other channel are not interchangeable; they cannot be considered as measures of the same phenomenon, but should be considered as evaluation of related but independent aspects (Sanavio, & Sica, 2004).

For instance, when clinician uses a test, he should consider it not as a real and unique description of a construct (e.g. depression) but as a description of only the subjective dimension of the construct, which is multidimensional (Sanavio, & Sica, 2004).

These three levels, or rather analysis' channels, form the "horizontal integration" of the assessment; "vertical integration", on the other hand, is given by the subsequent levels of in-depth test analysis it is carried out according to the psychologist's logical point of view, considering the interview, the observations, and then the more suitable measuring instruments (Sanavio, Bertolotti, Michielin, Vidotto, & Zotti, 2008).

In particular, the main tools available to carry out an assessment are classifiable into four categories (Groth-Marnat, 2009; Serra, Spoto, Ghisi, & Vidotto, 2015a): clinical

interview with observation, psychophysiological measurements, structured/semi-structured interviews, and self-report questionnaires.

1.1. Clinical interview and Observation

The first aim of clinical interview is the examination of the problem; it is a mainly hypothetical-deductive process (Sanavio, 2007). Clinical interview and observation provide a large amount of information, can follow adaptive logic, and allow taking advantage from multiple channels (verbal and non-verbal). Nevertheless, they require a great amount of time to be completed. Moreover, they are not always systematic and some inference problems may be introduced by the clinicians' bias, which could lead to wrong diagnosis and consequently ineffective treatment (Serra, Spoto, & Vidotto, 2015b; Spoto, 2011). The exclusive use of the clinical interview does not allow an exhaustive understanding of the individual's problems. In fact, it is known that patients often feel less uncomfortable in reporting their symptoms in self-report questionnaires rather than during interview. Moreover, during the clinical interview, the trust that the patient has in the clinician, as well as the ability of the individual to describe his symptoms, play a fundamental role. With regard to the observation, the expert must be able to focus on the object to be observed (e.g. a specific attitude) and, at the same time, to consider the context. For example, when the clinician is facing a depressed patient, it is necessary not only to consider patient complaints but also to understand through his behaviour his symptoms. In fact, the patient in the first interview can describe an inhibited depression and an agitated depression identically, but non-verbal clinical manifestations can play a crucial role in the assessment. Psychomotor agitation often manifests with incapacity to sit still, repetitive activity, pacing up and down, such as wringing of hands or even biting nails and/or lips (Goodwin & Jamison, 2007; Koukopuloulos et al., 2007). The ability of the clinician as well as the use of

observation grids for non-verbal behaviour can be of great help in observing the behaviour (Spoto, 2011).

1.2. Psychophysiological measurement

Psychophysiological measurement provides objective data, and deeply assess aspects that cannot be evaluated in other ways. However, it is limited in the areas of application, sometimes it has difficult accessibility, and it can be affected by artefacts. Moreover, the mood of the individual and the context in which he is can modify the results; For example, conditions that make the person uncomfortable may produce electrophysiological changes in heart rate or conductance beyond the experiment.

1.3. Structured interviews and Semi-structured interviews

Structured interviews (Van Zaane, van den Berg, Draisma, Nolen, Van den Brink, 2012) and Semi-structured interviews (Ferentinos, Paparrigopoulos, Rentzos, Zouvelou, Alexakis, Evdokimidis, 2011; Zimmerman, McGlinchey, Chelminski, & Young, 2012) are two tools used by the clinician at various stages of the assessment following a different procedure. Structured interviews are similar to orally administered questionnaires and follow a predetermined order of questions; they are standardized and, provide a numerical score on the analysed construct. Semi-structured interviews are of great importance. They do not follow a predefined sequence of questions but the various questions are investigated by the clinician during the interview taking into account any previous answer; so this is a clinician-led interview to find out the problem. The international reference point of semi-structured interview is the Structured Clinical Interview for DSM-IV (SCID). There are two versions of SCID: the first one refers to the axis I disorders of the DSM IV (SCID-I; First, Spitzer, Gibbon, & Williams, 1996), while the second one evaluate the axis II disorders of the DSM IV (SCID-II; First, Benjamin, Gibbon, Spitzer, & Williams, 1997). Semi-structured interviews, in our

perspective, are really interesting since they introduce the crucial concept of adaptivity in a systematic form. The selection of the areas to be investigated are guided by the clinician depending on the individual's responses; this means that not all questions are proposed to the patient, but only those in line with his symptomatology. The cons of this kind of interview are the great amount of time they require, and the possible inference errors introduced by clinicians. As in clinical interview and observation, even in semi-structured interviews, the clinician can make wrong logical inferences and consequently create problems in subsequent questions asked by the patient; In addition, the patient may become confused when answering, as he needs to respond immediately. In such cases, the result of the evaluation may not be reliable (Zimmerman et al., 2012).

1.4. *Self-report questionnaires*

Self-report questionnaires are often used in a hierarchical sequence; in the initial stages are explored several potentially problematic areas. These tests carry out a "broad spectrum" analysis and can be of great help to the psychologist for a first picture of the problem and for the focus of his subsequent examination strategies. Subsequently, in the later stages, more specific tests are used depending on the clinical case and its features noted in the previous stage. These tests offer the analysis of more specific and targeted constructs (Sanavio, & Sica, 2004). Therefore, the purpose of the tests is not just to classify the individual in a diagnostic category, but to provide the clinician with a more in-depth knowledge of the patient's problem. Tests should be considered as tools that make the work of the psychologist more detailed, allowing a more systematic and easy, clear and rapid exploration (Sanavio, & Sica, 2004). Tests allow to systematically and quickly collecting a lot of information; moreover, they avoid the possible embarrassment of patients. Nonetheless, they redundantly (non-adaptively) investigate constructs and they are conceived to provide a quantitative numeric score that does not

account for the qualitative information collectable from the single answer of the individual (Serra, Spoto, & Vidotto, 2015b).

2. Cognitive Behavioural Assessment (CBA 2.0)

According to Sanavio and Sica, "psycho-diagnostic examination is not a passive collection of information, but an active process, essentially similar to a process of problem solving and decision-making: a complex process of collection and processing information about the individual case" (Sanavio & Sica, 2004, p.9). A multidimensional Assessment integrates information and measurements from different levels.

Clinicians should not only consider the data that supports their hypotheses using only one channel, but also carefully consider or even list evidence that may not support their hypotheses take into account all the information's levels. This will likely reduce bias (Sanavio, 2007; Groth-Marnat, 2009).

In the Italian context, the Cognitive Behavioural Assessment test battery (CBA 2.0; Sanavio, Bertolotti, Michielin, Vidotto, & Zotti, 1997) was one of the first attempts aimed to mimic the clinician's evaluation including both horizontal and vertical integration.

This tool allows:

- To collect data;

- To analyse clinical indexes (included also the reliability of the results)

- To produce a preliminary report on the results

- To advise on areas which should be investigated further through the score obtained in the various primary scales and the analysis of so-called "critical items" (i.e. suicidal behaviours)

- To select new tests in order to understand specific clinical features (secondary scales)

- To indicate treatment strategies for a single case.

The authors of this self-evaluation battery wanted to develop a broad-spectrum instrument for clinical and therapeutic practice. CBA 2.0 suggested an attempt for a comprehensive approach to the psycho-diagnosis; the underlying procedure to analyse the individual replies is founded on the idea of a continuous progress in the process of diagnosis in order to decrease the uncertainties and strengthen the assumptions on the variables involved in the evaluation process (Bertolotti, Zotti, Michielin, Vidotto, & Sanavio, 1990). Authors consider the assessment in a multidimensional perspective; the information collected by self-report tools are integrated with the information obtained from other approaches (i.e. horizontal integration). CBA 2.0 allows to achieve the vertical integration thought the primary and secondary scales. The primary scales have the role of giving a first picture of patient's problems; they are broad-spectrum tests to explore several potentially symptomatic areas. The primary scales consist in ten sections. Sheet 1 collects personal data, while Sheet 4 evaluates personal life events (i.e. the educational and school situation, the affective relationships, the general state of health, eating and sleep habits, motivation to psychological treatment etc.). These two sections are useful for providing general relevant information to the clinician (such as interviews); they allow getting a more complete view of the individual case (horizontal integration). Sheet 2 is the State-Trait Anxiety Inventory X1 (STAI-X1; Lazzari & Pancheri, 1980; Spielberger, Gorsuch, & Lushene, 1970) and consider the state-anxiety of the individual when he/she starts to complete the test; it is mainly a measure of reliability of the test as a high level of anxiety can alter the performance of the subject. Sheet 3 (STAI-X2) and Sheet 10 (STAI-X3) assess the individual's anxiety (respectively trait-anxiety and state-anxiety) by means of the test STAI-X (Lazzari & Pancheri, 1980; Spielberger, Gorsuch, & Lushene, 1970). The sheet 5 is the brief form of the Eysenck Personality Questionnaire (EPQ/R; Eysenck, & Eysenk, 1975) and it

evaluate stable dimensions of personality such as the introversion-extroversion, emotional stability, anti-social behavioural etc. It consist of four main scales N (Neuroticism), E (Extroversion), P (Psychotic) and scale L (Lie); the latter one aims to provide a measure of the propensity of the subject to give a falsely "positive" profile. The sheet 6 is Psycho-Physiological Questionnaire brief form (QPF-R; Sanavio Bertolotti, Michielin, Vidotto, & Zotti , 1986) and it provides an evaluation of stress and psychophysiological symptoms. Sheet 7 is the Fears Inventory (IP-R; Sanavio et al, 1986; Wolpe & Lang, 1964); it evaluates individual's fears and the relationships between problem situations and emotional reactions. Sheet 8 is the Depression Questionnaire (QD; Sanavio et al., 1986) and assesses depressive symptoms. Finally, the sheet 9, Maudsley Obsessional-Compulsive Questionnaire brief form (MOCQ-R; Hodgson & Rachman, 1977; Sanavio & Vidotto, 1985) consists of three sub-scales: Checking, Cleaning and Doubting-Ruminating, which investigates the three sub-dimensions of the Obsessive Compulsive Disorder. All of these 10 scales form the first broad-spectrum evaluation. In line with the vertical integration of assessment, the secondary scales consist in much more specific tests, which should be used deepening on the clinical evidences obtained in the previous phase; they provide analyse of specific and target constructs. If a patient obtained a high score to QD, specific tools to relieve the severity of depression (i.e. Beck Depression Inventory) will investigate the depressive symptoms.

One of the main aims of CBA 2.0 project was to create a pathway as faithful as possible to the logical process of human operator (i.e. in circumscribing problems to successive phases and proposing hypotheses concerning therapy). The CBA 2.0 provides a descriptive computerized report of the patient score. It includes: analysis of the reliability of the test, high scores obtained in potential problem areas and finally positive responses to critical items. Although this attempt represented a great

improvement when compared with other similar researches developed in that period, CBA 2.0 still presents many limits when compared with what clinician needs (Serra et al., 2015b).

One of the most important critical things is that the question-answer progression followed a mechanic deterministic order. There were no ideas about how to simulate a well-articulated interview, which could be effective, efficient, and considering the natural and logical flow of question and answers contents.

A second critical aspect was related to the unsatisfactory use of the resulting scores. When a test is administrated, indeed, different combinations of symptoms may produce the same score in a way that such information may not be quite useful in clinical practice (i.e. the answer to each single item is not taken into account). Individuals may be classified under the same diagnostic category if they meet sufficient criteria even when they present dissimilar clinical features and are leaded by different underlying psychological mechanisms (Serra et al., 2015b).

A new formal approach, which gives a new perspective to the CBA 2.0 in terms of automatic-assisted procedures, is designed to cope with these problems. Formal Psychological Assessment (FPA; Spoto, Stefanutti, & Vidotto, 2010; Spoto, Bottesi, Sanavio, & Vidotto, 2013) will be discussed in the following chapter. In this work, FPA aims to overcome both the problem of adaptability and to be able to go beyond the numerical score in the assessment of mood disorders.

3. Adaptive Assessment

Psychological assessment has been based primarily on subjective judgment of clinician and classical psychometric test theory. Despite all the pros of the clinical interview, the problem of the subjective inferences of the clinician can cause errors in later evaluation

and diagnosis (Nordgaard, Sass, & Parnas, 2013). Indeed, the clinician's evaluation could be affected by underestimation or overestimation of patient's symptoms. Regarding the psychometric approach, typically, a total score determines the impairment level, which requires that all the same items are administered to all respondents. This last approach is primarily data oriented, and the product is often a series of scores. The score's descriptions are typically unrelated to the person's overall context and do not address unique problems the person may be facing (Hayes, Nelson, & Jarrett, 1987). In contrast, psychological assessment attempts to evaluate individual data in a broad perspective, with its focus being individual problem solving and decision-making.

Psychological assessment should include the evaluation of individual specific features. The central role of the clinician performing psychological assessment is that of an expert in human behaviour who must deal with complex processes and understand test scores in the context of a person's life (Groth-Marnat, 2009).

Thus, rather than just knowing the labels and definitions for various types of anxiety or thought disorders, clinicians should also have in-depth operational criteria for them. For example, the construct of depression, as represented by the score, can sometimes seem misleadingly straightforward. Depression can manifest with a variety of different symptoms that may be due to a different culture or a different aetiology (as reported in the Chapter on mood disorders). Only through personalized assessment can be possible to distinguish these conditions (Groth-Marnat, 2009). Unless clinicians are familiar with these areas, they are not adequately prepared to understand different types of depression (i.e., agitated depression, depression with flight of ideas, inhibition of depression, etc.).

An alternative to administration of a full scale achieving a personalize assessment is adaptive testing. It means that each individual may receive different scale items that are targeted to their specific impairment level (Fliege, Becker, Walter, Bjorner, Klapp, &

Rose, 2005). In adaptive testing, a person's initial item responses are used to determine a provisional estimate of his or her standing on the measured trait (for example, depression or anxiety) to be used for the selection of subsequent items (Wainer, 2000). This form of testing has recently emerged in the field of knowledge and mental health research (Falmagne & Doignon, 2011; Weiss, 2004). Procedures based on item response theory (Embretson & Reise, 2013) can be used to obtain estimates for individuals (for example, severity of depression) to more efficiently identify suitable subsets of item for each individual (Gibbons et al., 2008). In particular in the last years several studies demonstrated that diagnostic instruments could benefit substantially from modern statistical approaches like models of item response theory (IRT), e.g., the Rasch model. Indeed, by using IRT-modelling it was shown that unidimensionality, an important aspect of test theory, cannot be taken for granted. For example, if a patient suffering from a severe somatic illness reported somatic symptoms in a depression questionnaire those symptoms may be ascribed to the somatic illness or a depressive episode (Forkmann et al., 2009). Moreover, it was shown that questionnaires could be shortened without loss of information. This testing approach is referred to as computerized adaptive testing (CAT) and can be applied to achieve a more effective assessment (Petersen, Groenvold, Aaronson, Fayers, Sprangers, & Bjorner, 2006). The main idea is to administer a small, optimal number of items to the individual without loss of measurement precision and according with his previous answer. This process mimics the semi-structured interview, with the difference that the inferences are made by an algorithm which considers all the information and step by step goes through the assessment following logically correct process (Spoto, 2011).

Eggen and Straetmans (2000) combined IRT with statistical procedures, like sequential probability ratio test and weighted maximum likelihood, for classifying people under exam. Other systems use Bayesian statistical techniques instead of IRT in the evaluation

of students' knowledge (e.g. EDUFORM, Nokelainen, Silander, Tirri, Nevgi, & Tirri, 2001; and PARES, Marinagi, Kaburlasos, & Tsoukalas, 2007).

In the field of knowledge assessment ALEKS (Assessment and LEarning in Knowledge Spaces) is a complex system able to adaptively assess a subject's knowledge and provide a consequent learning individualized path (Grayce, 2013; Donadello, Spoto, Sambo, Badaloni, Granziol, & Vidotto, 2016). Starting from a set of items on a specific topic, the output of ALEKS system is the subset of items, which the subject is able to reply; this subset is called "knowledge state" and it refers to the level of knowledge of the individual in a particular field.

However, the formulation of the adaptive algorithm is even more difficult in the clinical setting. In fact, the objectivity of the questions and therefore of the answers given by the subject is much more questionable, and the probabilities of making mistakes in the answer increase. Despite this, research has demonstrated that both item response theory and CAT (Baek, 1997) can be applied to the measurement of attitudes and personality variables (Reise & Waller, 1990). In the clinical context, Spiegel and Nenh (2004) developed an expert system, which calculates possible symptom combinations and returns all possible risk diagnoses. Yong and colleagues (2007) developed an interactive self-help system for depression diagnosis that provides advice about patients' levels of impairment. Simms, Goldberg, Roberts, Watson, Welte, & Rotterman (2011) developed the CAT for Personality Disorders (CAT-PD) aimed at realizing a computerized adaptive assessment system. CAT has been applied also in developing adaptive classification tests by means of stochastic curtailment using CES-D for depression (Finkelman, Smits, Kim, & Riley, 2012; Smits, Finkelman, & Kelderman, 2016).

Gibbons and colleagues (2008) used the combination of item response theory and computerized adaptive testing (CAT) in mood and anxiety disorder assessment. In particular they applied a bifactor structure, consisting of a primary dimension and four

sub-factors (mood, panic-agoraphobia, obsessive-compulsive, and social phobia). Participants completed the Mood and Anxiety Spectrum Scales (MASS) at two times. The first administration was used to define an adapting testing version of the MASS, the second confirmed the functioning of CAT in live computerized testing. Authors created item banks with a large item pool, and being able to administer a small set of the items most relevant for a given individual with no loss of information, allowing a strong time reduction and consequent patient and clinician burden. A chart review was performed for six patients with mood disorders (three major depressive disorder and three bipolar disorder) who were interviewed by the psychiatrist. Most of the CAT items that were endorsed positive were not documented in the six patients' psychiatric evaluation through SCID-I. These items included clinically important information, such as a history of manic symptoms, potentially risky behaviours etc. This last study is an important example of how adaptive testing can be effective. Despite this, it has several limitations: first, the proposed model is totally deterministic; it starts from a theory based on the factorial structure and does not take into consideration the possibility that the subject's answers are not corrected. A second limitation, according to the bifactor model, there is only one main dimension and the sub-dimensions related; so, if this condition is not satisfied the model can not be used. Finally, this model works only if each item loads on a primary dimension and no more than one sub-dimensions. If items are related to multiple sub-dimensions, they will not be appropriate for the bifactor model and therefore CAT is not applicable.

However, although there have been several attempts to apply adaptive clinical assessment, as far as we know, no system was able to combine adaptability, quantitative and qualitative information, and estimate error parameters through a probabilistic model.

The Formal Psychological Assessment, and its application to mood disorders, is the core of this work, and represents a further step to overcome the obstacles encountered up to now in adaptive testing.

3. Conclusion

The assessment is a crucial moment in the therapeutic process; the clinicians, after collecting as much information as possible on the patient, must formulate diagnostic hypotheses in a short period to plan clinical interventions. The quality of clinical evaluation is crucial for both diagnosis and treatment; In fact, a misdiagnosed psycho-diagnostic evaluation could result in therapeutic failures, dissatisfaction and patient suffering.

Keeping in mind the previous background, we can list the key features of an ideal assessment tool:

1. Adaptive logic as a semi-structured interview that allows examining in depth only the individual's symptomatic areas (questions are guided by the psychologist depending on patient answers to previous questions).

2. Ability to perform systematically correct inferences, avoiding inference errors by clinicians.

3. Exhaustivity similar to a clinical interview; to get all the information needed to contextualize the problem of the single case. The tool should be completed for the investigation of the target construct.

4. Rapid administration (as a test), allowing to collect a lot of important information in a short time.

5. Ability to provide quantitative and qualitative information. The tool should be able to have a numerical score that allows to classify the individual into a diagnostic category (nomothetic approach); on the other hand it should underline

the set of peculiar symptoms of the individual case in order to distinguish its symptomatology and to treat it properly (ideographic approach).

6. Validity and Reliability; the instrument must be built following all the psychometric criteria, tested and validated before use.

To achieve this result, the ideal tool should have the positive features of the interview and the strengths of the tests without the critical aspects of both. In the next section, the new methodological proposal will be presented and discussed.

CHAPTER 3:

The Formal Psychological Assessment

1. Introduction

The FPA is a methodology potentially capable of maximizing the advantages of both semi-structured interviews and self-report questionnaires by overcoming the limitations of these tools and managing the problems of traditional assessment (Spoto, Bottesi, Sanavio, & Vidotto, 2013).

The ability to analyse clinical symptoms is important when evaluating the responses to a questionnaire. FPA is able to go beyond the score of the patient and investigates the diagnostic features implicated by the responses (Serra et al., 2015a). The crucial underpinning that represents the starting point of FPA is consideration of all the information that can be collected from a patient's numeric score on a questionnaire. For instance, if a nine dichotomous items scale is administered to a patient and the clinical cut-off score of the scale is 7, there are 46 different clinically significant response patterns (one pattern with score 9, nine patterns with score 8, and 36 possible patterns with score 7). It is clear how each of these patterns may convey clinically different information about the patient' symptoms. Notice that all this information is already included in the questionnaire, even if the mere score somehow hides it.

Nevertheless, at the present time, the only ways the clinician has to account for the specific information endorsed by the pattern are:

    A. To read all the items the patient has answered affirmatively, and from them, try to deduce his/her clinical situation (it is noteworthy that this solution is applicable only when the questionnaire counts a low number of items, and that

this operation cannot be carried out when tolls like the MMPI-2 are administered).

B. To further, investigate this issue through psychological interview. Clinical interview, although essential during assessment, if used alone may not be reliable. As demonstrated by several studies, the clinician may overestimate some symptoms and underestimate others, or make logical inferences that are not entirely correct.

Both of these solutions do not provide any standardized procedure comparable to the systematic scoring of the questionnaire. The FPA aims to provide an in depth analysis of the specific response pattern observed, thereby informing the clinician about the actual diagnostic configuration of the patient at hand (Spoto et al., 2013). This opportunity is assured by an a priori analysis of the clinical elements investigated by each of the items of the questionnaire. Such analysis is the deterministic skeleton on which it is possible to implement a probabilistic adaptive procedure capable of mimicking a semi-structured interview within the frame of a questionnaire. By highlighting the specific clinical elements investigated by each single item of a questionnaire, FPA highlights the differences among patients that would otherwise be hidden by the simple score. From a clinical perspective, it allows for an idiographic and nomothetic diagnosis (Serra et al., 2015a). Moreover, this approach, compared to both the IRT and the classical psychometric approach, allows: first, a higher level of reliability and validity of the measurement; second, an ability to process in a faster way a higher number of information in the vertical integration inference process through adaptive algorithm (Spoto, 2011).

The FPA (Spoto et al., 2010; Spoto et al., 2013) developed from the conjunction of two mathematical psychology theories: Knowledge Space Theory (KST; Doignon and

Falmagne, 1985, 1999; Falmagne and Doignon, 2011) and Formal Concept Analysis (FCA; Wille, 1982; Ganter and Wille, 1999).

2. Knowledge Space Theory

Knowledge Space Theory (KST) was born from a search of an efficient and objective measurement of the knowledge acquired by an individual in a learning program. It is applied in the field of education, and it consists of a computerized procedure for assessing the knowledge of the individual in an adaptive way. In the KST language, the items proposed to the subject are sorted according to their difficulty; in this way, the teacher will be informed about what the student already knows and what he/she is ready to learn in the following steps. According to KST given a set of items on a specific topic, the output of the assessment (called "knowledge state") is the subset of items that the individual under evaluation has showed to master. The basic concepts of this theory are the knowledge domain (Q), the knowledge state (K ⊆ Q), the knowledge structure ($\mathcal{K}$), the knowledge space, the skill-map (Q, S, f), and the probabilistic knowledge structure (P, K, π). These elements provide the basis for building the FPA model.

- ✓ *Knowledge domain* (Q): is the set of questions that can be asked about a topic in order to investigate an individual's knowledge.
- ✓ *Knowledge state* (K): is the subset of Q containing all the questions that an individual is able to answer correctly.
- ✓ *Knowledge structure* ($\mathcal{K}$): is a collection of knowledge states (subset of Q) which contains at least the empty set (Ø) and the total set Q. In the traditional formal notation a knowledge structure is denoted as (Q, $\mathcal{K}$) where Q represents the knowledge domain and $\mathcal{K}$ represents the collection of subsets included in the

structure. The knowledge structure is a representation of the implications among the items belonging to Q.

Given a knowledge domain Q = {a, b, c}, a, b, c are the questions to evaluate. A possible knowledge structure could be $\mathcal{K}$= {Ø, {a}, {c}, {a, b} {a, b, c}}, where Ø, {a}, {c}, {a, b}, {a, b, c} are the different states of knowledge K_0, K_1, K_2, K_3, K_4.

It is important to note that in this precise knowledge structure the mastery of the item *a* is a prerequisite for the mastery of item *b*; indeed, there is no knowledge state in $\mathcal{K}$ that contains *b* and does not contain *a*. If a subject does not solve the item *a*, he will not solve the question *b* (except in the case of lucky guess- i.e. false positive -; η). In KST this means that *a* is a prerequisite for *b*.

- ✓ *Knowledge space*: is a particular class of knowledge structures. It is defined as a knowledge structure in which the union of any group of knowledge states generates a new subset already included in the knowledge structure (that property is defined as closure under union). An interesting property of a knowledge space is that more than a single set of prerequisites are allowed for an item. This means that the same item can be solved using different solution strategies.

- ✓ *Skill-map* (Q, S, f): The skill map is defined as a triple (Q, S, f) where Q is a non-empty set of items, S is a non-empty set of skills, and *f* is a mapping from Q to 2^s \ {Ø} (i.e. the power-set of S excluding the empty-set; Doignon & Falmagne, 1999). The skill map concept is important for delineating a knowledge structure. For this purpose, three models are used: The Disjunctive Model, the Conjunctive Model and the Competency Model (Doignon & Falmagne, 1999). In this work, only the first two will be considered, translated into the language of the Formal Psychological Assessment. In the disjunctive model, in order to master an item it is sufficient to have at least one of the

required skills, on the other hand, in the conjunctive model, all the skills are necessary in order to achieve a specific item. The knowledge structures corresponding to these two models are respectively closed under union (knowledge space) or intersection (each intersection of sets of states is included in the structure). The skill-map can be represented by a Boolean matrix with the items in the rows and the skills in the columns.

- ✓ *Probabilistic knowledge structure* (P, K, π): Since the formal deterministic model does not fully reflect reality, it is necessary to define a probabilistic model. The Basic Local Independence Model (BLIM) allows to assign probability values to the different states of a knowledge structure; responses to each item are considered locally independent. This model has been applied in a number of different contexts (e.g. Falmagne, Koppen, Villano, Doignon, & Johannesen, 1990; Spoto, Stefanutti, & Vidotto, 2010). In the model, starting from the probabilistic structure (Q, K, π) and a response pattern R ⊆ Q, a probability distribution $p(R, K)$ can be derived for each response pattern R given a knowledge state K, in the clinical structure. Formally:

$$p(R) = \sum_{K \in \mathcal{K}} \rho(R, K) \pi(K)$$

In the probability model, the response function is also determined on the basis of the two error parameters: the *lucky guess* and *careless error* (also called false positive and false negative errors). In case of a careless error (β) a subject does not solve an item that he is able to solve, while in case of a lucky guess (η) the subject solves an item that he is not able to solve.

Formally:

$$\rho(R, K) = \left[\prod_{q \in K \setminus R} \beta_q\right]\left[\prod_{q \in K \cap R}(1 - \beta_q)\right]\left[\prod_{q \in R \setminus K} \eta_q\right]\left[\prod_{q \in \overline{K \cup R}}(1 - \eta_q)\right]$$

All KST described concepts will be used in the application of FPA, so they will be translated into a clinical language suitable for psycho-diagnostic evaluation.

3. *Formal Concept Analysis*

Formal Concept Analysis (FCA) is based on a simple observation: when we think to an object, we identify the characteristics that define it and that allow distinguishing it from other objects; on the other hand, given a set of features, we are able to identify which object it describes. The main concepts of FCA are the *formal context* (G, M, I), and the *formal concept* (A, B).

The *Formal context* is defined as a triple (G, M, I). G is a set of objects, M is a set of attributes and I is a binary relation between the set of objects and the set of attributes. Starting from a Boolean Matrix objects are placed in the rows, while the attributes are placed in the columns. In the formal context, each row represents an object and each column represents an attribute. Whenever in the cell is present value 1 means the relation *gIm* holds; in other words it means that the object *g* has the attribute *m*. Between the objects and the attributes of a formal context a *Galois connection* is defined. For all the sets A $\subseteq$ G and B $\subseteq$ M, the following two transformations define the Galois connection:

$$A' \coloneqq \{m \in M | gIm \; \forall g \in A\}$$

$$B' \coloneqq \{g \in G | gIm \; \forall m \in B\}$$

It means that A' is the collection of all the attributes that all the objects in A have in common. Dually B' is the collection of all the objects that possess all the attributes in B.

It is now possible to introduce a fundamental notion of FCA. The pair (A, B) is called a *formal concept* if it satisfies the following two conditions: $A = B'$ and $B = A'$. The so called *extent* A of the formal concept contains exactly those objects of G that have all the attributes in B; the so called *intent* B of the formal concept includes exactly those attributes satisfied by all the objects in A. A sub-concept super-concept relation is then defined in the following way:

$(A_1, B_1) \leq (A_2, B_2) \leftrightarrow A_1 \subseteq A_2$

or equivalently:

$(A_1, B_1) \leq (A_2, B_2) \leftrightarrow B1 \supseteq B2$

A concept is of a lower level when it has a larger extent (or equivalently a smaller intent). The concepts of a context form a *complete lattice* (Birkhoff, 1937, 1967) that is called the *concept lattice* of (G, M, I). The intents of a concept lattice are closed under intersection (i.e. each intersection of sets of attributes is included in the lattice). The collection of the complements of the intents of a formal context is closed under set-union and then it is a knowledge space (Rusch & Wille, 1996).

Thus, the *concept lattice* can be described as a particular diagram, which describes the order relationships among the objects within a knowledge space. A lower order relationship is defined when a set of objects in a formal concept is contained in another formal concept. From this definition, we can deduce that larger is the set of attributes, the smaller is the set of objects characterized by those attributes. By increasing the number of attributes, the object can be specified. FCA concepts are integrated with those of KST in the application of the *Formal Psychological Assessment.*

4. *Formal Psychological Assessment description*

FPA approach, by the conjunction of the two described theories, allows the development of an instrument with multiples benefits based on the formal representation of the relationship between the items of a questionnaire and a given set of clinical criteria. The Formal Psychological Assessment can be applied both for the construction of new evaluation tools and for the description of the self-report instruments used in clinical practice.

In the first case, FPA allows the creation of an efficient tool:

- ✓ *Adaptive,* where the item proposed to the patients depend on his/her previous answer in a process that mimics semi-structured interviews.

- ✓ Able to provide *quantitative and qualitative information.* It does not only provide a quantitative numeric score, but it also explores the entire patient's symptoms configuration.

In the second case, FPA allows to analyse in detail the self-report questionnaires in order to underline both the strengths and the weakness.

In FPA, each item included in a clinical self-report questionnaire (or interview) is defined as an *object*. Each object can be described on the basis of a set of elements referring to a given theoretical framework. Such elements (which can be either clinical symptoms or the decomposition of the diagnostic criteria used to specify one or more clinical disorder), are named *attributes*. Thus, each object can be related to the set of attributes it endorses. For instance, in characterizing the items (objects) of a given clinical self-report questionnaire, the attributes may be represented by the DSM-5's diagnostic criteria of disorder the questionnaire is supposed to investigate.

Theoretical flexibility is one of the major strengths of the FPA. In fact, the same objects can be described in terms of attributes by referring to different appropriate frameworks (Spoto et al., 2013). Each item may investigate one or more attributes and each attribute

can characterize one or more items. For example, the item "I am less interested in sex than I used to be," related to depression, investigates the attributes "Diminished interest and pleasure" and "Decreased interest in sex," which represent two diagnostic criteria for depression in the DSM-5. On the other hand, the attribute "Diminished interest and pleasure" is investigated by several items (i.e., "I have lost most of the interest in other people or things", "I am less interested in sex than I used to be", "I do not want to do anything" and "I seem to have lost interest in the future").

Starting from a set of objects (items) and a set of attributes (clinical criteria), a Boolean matrix can be built assigning to each object its own set of attributes. The items are placed in the rows of the matrix, and the attributes are placed in the columns. Every time an item investigates a specific attribute, the corresponding cell of the matrix will contain "1", otherwise the cell will contain "0." In FPA this matrix represents the *clinical context*. The entire set of objects is the *clinical domain* of the clinical context. The *clinical state* of a patient consists of the subset of items he/she answered affirmatively. It is noteworthy how each clinical state (depicted by the response pattern endorsed by a patient) correspond to a subset of attributes. Thus, even if two patients respond affirmatively to the same number of items (i.e., obtain the same score to the questionnaire), the representation of their two states in terms of attributes are systematically different, if the items affirmatively answered are also different. Thus, different states may have the same clinical score (Serra et al., 2015a), but will collect different attributes. This is fundamental from a clinical point of view, since it allows for the analysing and, therefore, the treating, of each subject individually, according to his/her symptoms configuration. The configuration of symptoms as an output also makes possible to overcome the problem of gender differences in depression assessment (Santor, Gregor, & Welch, 2006); indeed, the clinician will not have only the numerical score, but the whole symptomatology of the individual. The clinical context is the

Boolean representation of the *clinical structure,* which is the set representation of the implications among the items of the domain. The clinical structure contains all the clinical states that are formally expressed by the matrix. In fact, not all the subsets of items are admissible response patterns given a theoretical framework (thus, given the formal context). For instance, if a given item *i* endorses attribute *a*, while item *j* endorses attributes *a* and *b*, the clinical state corresponding to {*i*} is admissible. On the other hand, the state {*j*} is not admissible since a person who affirmatively answers item *j* is supposed to present both attributes *a* and *b*, and thus, he/she should affirmatively answer even item *i*. For this reason, the state {*i,j*} is admissible too. In this case, item *i* is said to be a prerequisite of item *j* since there is no state in the structure that includes the latter but not the former. The prerequisite relation among the items, obtained from the matrix through the formal mathematical passages, can be represented as a *complete lattice* depicting the clinical structure. Thus, the implications of the items form prerequisite relations can be described in the same way as KST but in a clinical context. In other words, the prerequisite is an item that contains the same attribute "a" of another item that contains the same attribute "a" and another one or more (for example "b"). The first item is needed to get a positive answer to the second item. The prerequisite relation among the items obtained from the matrix can lead to the development of adaptive and qualitative tools as well as quantitative tools (Donadello et al., 2016; Spoto et al., 2010).

From a strictly methodological perspective, the clinical structure derived from the matrix can be developed in various ways using the concepts of KST and FCA. From a clinical perspective, we would like to explain two model derived from the clinical structure: The conjunctive model and the disjunctive model, which have been mentioned for the KST.

➢ *Conjunctive model* assumes that if an individual positively responds to an item, he/she should endorse all the attributes, in FPA clinical criteria (i.e. symptoms) investigated by that item. With this model, the affirmative answers are more informative than the negative ones; indeed each affirmative answer inform us that the patients has all the symptoms investigated by that item. For example, item j "I do not have need and energy to have sex" could include two symptoms. The two clinical criteria associated are "decreased interest in sex" and "fatigue or energy loss" for the major depressive episode of DSM-5. The conjunctive model assumes that if the patient answers "yes" to the item j, then he/she will have both symptoms (attributes) associated with that item.

➢ *Disjunctive model* assumes that if an individual responds positively to an item, it means that he/she has at least one attribute investigated by that item. In this model, the negative answers are the more informative because they inform us that the patient has none of the symptoms signalled by that item. In other word, in the item j "I do not have need and energy to have sex", if the patient replies "yes" it can mean three possible thinks: A) the patient has the diagnostic criteria "fatigue or energy loss" and not "decreased interest in sex". B) The patient has the diagnostic criteria "decreased interest in sex" and not "fatigue or energy loss". C) The patient has both the clinical criteria. Otherwise, if the patient replies "no" it means that he/she has not these two clinical criteria.

From the clinical point of view, we prefer to use the conjunctive model as each affirmative answer inform us that the patients has all the symptoms investigated by that item. Then in our model, we assume that when the individual replies "yes" he has all the criteria endorsed by the item in question.

To sum up, the first step of FPA methodology is the deterministic model construction, which consists of the construction of the matrix assigning to each item of the scale the subset of attributes it investigates.

The second step concerns the construction of the clinical structure from the attributes assignment. The result can be represented as a lattice where each node represents a clinical state and its set of attributes (Spoto et al., 2010). The lattice is a deterministic representation of the prerequisite relation among the items of the domain.

It is evident how a completely deterministic approach is inadequate for assessment in clinical practice for three main reasons:

- ✓ First, not all clinical states have the same probability of occurring.

- ✓ Second, in self-report tools, problems with patient insight or with item wording may prevent a perfect correspondence between the observed response pattern and the actual clinical condition of the patient.

- ✓ Third, the deterministic model needs to be tested on real data.

Therefore, a probabilistic approach is needed and, it is applied in the same way of KST. The basic local independence model (BLIM; Doignon, & Falmagne, 1999) is a probabilistic model that defines a probabilistic clinical structure where a probability value is assigned to each clinical state. In the BLIM, the responses to each item are locally independent given the clinical state of a subject. Starting from the probabilistic structure, the probability of a response pattern depends on the conditional probability of that pattern given an underlying clinical state (for each state; Doignon & Falmagne, 1985). The false negative (β) and the false positive (η) rates for each item define the conditional probability (Falmagne & Doignon, 2011). The clinical structure, by means of the probabilistic weights obtained through the application of the BLIM, could be used to implement an algorithm.

The last step in the application of FPA is the implementation of the obtained structure, weighted through the estimated parameters, into an adaptive algorithm for the clinical assessment. Such an algorithm can be roughly divided into three main parts corresponding to different moments of inferential process. First, the algorithm has to select, among all the possible items to be asked, the most informative one. To perform this task, one of the most reasonable solutions is to identify the item that best splits the probability mass. Later the system registers the reply provided by the individual and it automatically updates the new probability of the states. Indeed, in case of positive reply, the system increases the probability of the states that include the asked item, and decreases the probability of the other states. If the reply of the individual is "no" the system increases the probability of the states which do not include the asked item. In this way, a sequence of question is asked and at the end, one of the all clinical state of the structure should achieve an high probability value (with a fixed cut-off). This clinical state represents the most likely symptomatic representation of the patient's situation regarding a specific disorder. At this point, the algorithm stops and provides the clinician with the score, the response pattern, and the attributes configuration (all the symptoms complained by patients).

In this way, FPA allows for an adaptive, quantitative and qualitative tool: adaptive because, based on the structure, it selects each question to maximize the collectable information; quantitative because it could provide a numerical score; and qualitative because it provides information about all the subjects' symptoms.

The next three chapters of this work will be devoted to describing how we applied the FPA in assessing mood disorders. This dissertation work is divided into four parts:

- The first will describe the application of the FPA method in the analysis of seven self-report questionnaires for the evaluation of the major depressive episode.

74

- The second will describe the construction and validation of a new self-report questionnaire using the FPA methodology that allows differentiating patients with the same score but different depressive symptoms.

- In the third, the application of the adaptive algorithm to the new questionnaire via the FPA will be shown.

- The last research deepens the topic of agitated depression, in a study of 3750 patients with mood disorders. This study, carries out in England (data come from BDRN of UK) underlines the need of new effective tool for the differential diagnosis in Major Depressive Episode.

CHAPTER 4

Application of FPA to analyse the relations between *item* and *diagnostic criteria* in Major Depressive Episode

1. Introduction & Research aim

The increase of depression in the last few years is a debated topic (Serra et al., 2015a).

Some authors argue that, nowadays, depressive disorders with bipolar disorder are the most common type of disease in the world, though often unrecognized and inadequately treated (Kupfer et al., 2012; Lancet Editorial, 2012).

The correct identification of depression during the assessment phase is a critical issue.

Despite this, many authors report that the evaluation tools available to the clinician are not effective enough for a proper identification of depressive symptoms for various reasons (Serra et al., 2015a).

In a critical study, Balsamo and Saggino (2007) underlined the strengths and weaknesses of some important self-assessment tools of depression. The purpose was to avoid confusion in clinical practice. Specifically, the study explored the psychometric properties of six self-report measures of depression. The Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996). The Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977). The Zung Self-Rating Depression Scale (Sakamoto, Kijima, Tomoda, & Kambara, 1998). The Clinical Depression Questionnaire (CDQ; Krug and Laughlin, 1976). The Questionnaire for Depression (QD), included in the Cognitive Behavioral Assessment 2.0 battery (CBA 2.0; Sanavio et al., 1986); and the D scale of the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway and McKinley, 1942). Balsamo and Saggino (2007) showed that

each scale reflects the authors' theories, which were constructed to measure different aspects of the same construct. Thus, each scale emphasized the evaluation of some symptoms and neglected the evaluation of others, leaving incomplete overall evaluation. Pettersson, Boström, Gustavsson, & Ekselius (2015) conducted a systematic review on evaluating depression tools, which revealed that only the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I; First et al., 1996), the Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998), and the Patient Health Questionnaire (PHQ-9; Manea, Gilbody, & McMillan, 2012) fulfilled the minimum criteria for sensitivity and specificity. Out of these three tools, only the PHQ-9 is a self-report measure that can be used for screening, diagnosis, monitoring, and measuring the severity of depression.

Regarding PHQ-9, although it is composed of nine items that correspond to the symptoms of depression according to DSM-IV, it does not distinguish insomnia from hypersomnia, and does not distinguish psychomotor retardation from psychomotor agitation. In fact, in PHQ-9 there is only one item for insomnia-hypersomnia and one item for psychomotor retardation-psychomotor agitation. Therefore, it is not potentially able to differentiate different depressive symptoms that, as such, should be treated in different ways.

The present work aims to describe a practical application of FPA to illustrate procedural issues, discuss the advantages of the approach, and show its potential for psychological assessment, relating to depression. In particular, in this work FPA is applied to analyze the "item content" of the most used self-report questionnaires for the depression's evaluation. Indeed FPA allows creating relations between "item content" and diagnostic criteria to the assessment of Major Depressive Episode (MDE). In keeping with previous research, the main task of the first study is to underline the strengths and the

weaknesses of widely used depression tools, in relation of their ability to investigate all the symptoms of MDE. As suggested by Balsamo and Saggino (2007) it is crucial to be aware of what aspects may not be investigated through the specific used assessment tool. To achieve this aim we use the main concept of FPA described in the Chapter above.

2. *Method*

2.1 Attributes' selection

In order to analyze the relations among a large set of items used to investigate depression through self-report measures and a set of symptoms of MDE, we mostly refer to three areas to derive the symptoms of Depression:

1. The Diagnostic and Statistical Manual of Mental Disorders (DSM-5; 2013). We have chosen to use the DSM-5 as it appears to be an exhaustive manual designed for both the researcher and the clinician (Reed et al., 2013)

2. The clinical features most frequently reported in literature to describe MDE.

3. Seligman's and Beck's etiopathogenetic theories.

As described in detail in the first chapter, depression is characterized by deep sadness and despair, hopelessness, helplessness, and worthlessness (Goodwin & Jamison, 2007; Otto, Andreas, Von Klitzing, Fuchs, & Klein, 2014). Furthermore, a depressed mood is associated with anhedonia (Gaillard et al., 2013; Goodwin & Jamison) apathy (Alexopoulos et al., 2013; Mulin et al., 2011) loss of motivation (Jormann, & Quinn, 2014) crying (Goodwin & Jamison; Koukopoulos et al., 2007), and irritability (Akiskal, & Benazzi, 2003; Henderson, Johnson, Vallejo, Katz, Wong, & Gabbay, 2013; Pedrelli et al., 2013). Feelings of guilt are frequent (Goodwin & Jamison, 2007; Singh, & Sharma, 2013) in more severe forms, they can result in delusion of guilt (Goodwin & Jamison, 2007). Sleep problems characterize depressed patients and frequently their

insomnia is "terminal" (waking early in the morning) or characterized by frequent nocturnal awakenings or by feelings of not being rested after waking up (Goodwin & Jamison, 2007; Hamoen, Redlich, & Weerd, 2014); conversely, some patients experience hypersomnia (Goodwin & Jamison, 2007). Work and social relationships are often severely compromised (Fried & Nesse, 2014; Goodwin & Jamison, 2007). Psychomotor retardation can present as simple motor slowing, but more often does so as ideation and speech slowing as well as concentration difficulties (Bracht et al., 2012; Goodwin & Jamison) which accompany fatigue and energy loss (Fava et al., 2014). Many patients experience agitation, which can manifest as restlessness, incapacity to sit still, torturing hands and/or hair or even biting nails and/or lips (Akiskal, & Benazzi, 2004; Goodwin & Jamison, 2007; Koukopoulus et al., 2007; etc.). Sexual disorders such as decreased libido can be observed (Goodwin & Jamison, 2007). Lastly, ideas of death such are usually associated with depressed mood (Goodwin & Jamison, 2007; Olin et al., 2012). Suicide is the most tragic consequence of depression and the number of suicides has not decreased since the use of antidepressants (Baldessarini et al., 2006a).

Beck's and Seligman's theories have been described at the end of the first chapter. Beck's model (Beck, 1991, 2005) categorizes typical beliefs and mistakes of depression as a cognitive triad that includes a negative view of self, a negative view of the world, and a negative view of the future.

Seligman's theory (Seligman, 1972), based on animal experimentation, suggests that depression is associated with the conviction that nothing can be done to face stressful life events. This is learned helplessness, which tends to be generalized to new situations with the expectation of having no control over the future (Abramson et al., 1978).

In the research we explored the symptoms derived from DSM-5 diagnostic criteria for major depressive disorder (15 attributes), Seligman's and Beck's theories (3 attributes) and, finally, attributes widely described in the literature (2 attributes):

✓ <u>apathy</u> (Alexopoulos et al., 2013; Mulin et al., 2011) as a state of indifference to the world, characterized by inability to express feelings and lack of will.

✓ irritability (Akiskal, & Benazzi, 2003; Henderson et al., 2013; Pedrelli et al., 2013) which is expressed with frequent spells of weeping, mood lability, nervousness, and marked reactivity.

Subsequently, tools and clinical symptoms of depression have been selected for the construction of the model according to FPA procedure. In line with FPA all the clinical criteria in Table 4.1 were placed in the columns of the Boolean matrix.

Table 4.1: the twenty attributes (clinical criteria) of the clinical context.

Attribute	Explanation
A1	Depressed mood
A2	Diminished interest and pleasure
A3	Decreased interest in sex
A4	Increase or loss of weight
A5	Gain or loss of appetite
A6	Insomnia or hypersomnia
A7	Agitation
A8	Psychomotor retardation
A9	Fatigue or energy loss
A10	Feelings of worthlessness (or Beck's negative view of self)
A11	Feelings of guilt
A12	Diminished ability to think and concentrate
A13	Indecision
A14	Recurrent thoughts of death
A15	Suicidal ideation or attempted suicide
A16	Beck's negative view of the world
A17	Beck's negative expectation of the future
A18	Seligman's learned helplessness
A19	Irritability
A20	Apathy

2.2. Object's description (Self-report Questionnaires)

Four self-evaluation questionnaires developed in English, one self-report questionnaire in French, and two self-report questionnaires in Italian were selected (each here presented in English).

- The Beck Depression Inventory II (BDI-II; Beck, Steer, & Brown, 1996) is one of the world's most widely used self-report questionnaires for the evaluation of depression. BDI-II contains 21 items that explore various facets of depression. Each item has four possible answers of increasing severity, for a total of 84 items (21x4). The recommender cut-off is 17. The tool appears to be both agile and sensitive.

- The Self-rating Depression Scale (SDS; Zung, 1965) assesses the level of depression. The tool explores affective, somatic, and psychological dimensions of depression. It consists of 20 items: two for the affective symptoms, eight for somatic symptoms and ten for cognitive symptoms. Items are evaluated on a 4-point scale that corresponds to: 1 = nothing or only for a short time; 2 = a little bit of time; 3 = a big part of the time; 4 = continuously or much of the time. The tool is very simple and quick.

- The Rome Depression Inventory (RDI; Pancheri, & Carilli, 1982) consists of a series of 25 items evaluated on a 4-level scale, from 1 (no) to 4 (very severe). The items of this tool use the phrases most frequently complained by depressed patients to describe their illness and discomfort.

- The Plutchik-Van Praag self-report depression scale (PVP; Plutchik, & Van Praag, 1987) was developed with 34 items to cover all the DSM-III diagnostic criteria for depression. Items are evaluated on a 3-point scale (0 to 2), where 0 = absent 1 = moderate; 2 = marked. Scores of 20-25 indicate a likely depressive

disorder. Since these diagnostic criteria have remained largely unchanged in DSM-5, this scale still holds great validity.

- The Carroll Rating Scale (CRS; Carrol, Feinberg, Smouse, Rawson, Greden, 1981), is a self-report version of the Hamilton Rating Scale for Depression (HAMD; Hamilton, 1960), consisting of 52 dichotomous items. Items can only have a "yes" or "no" answer.

- The Self-Assessment Scale for Depression (SAD; Cassano & Castrogiovanni, 1982). It consists of 31 items. Items are evaluated on a 4-level scale, from 1 (absence of symptom) to 4 (maximum severity). The authors tried to use a language close to that of patients in the formulation of the items to contribute to a better comprehension of questions and a higher reliability of the instrument.

- Finally, the Center for Epidemiological Studies Depression (CES-D; Radloff, 1977; Eaton, Muntaner, Smith, Tien, Ybarra, 2004) has been one important instrument in depression epidemiology since its first use in Community Mental Health Assessment Surveys in the 1970s. The self-report version is widely used and consists of 20 items.

In conclusion, the total number of items adds up to 266. All these items were placed in the rows of the Boolean matrix.

2.3. Procedure

Every item in the clinical self-report questionnaires described above was initially considered.

These items became the objects of the matrix and represented the rows of the matrix for an initial 266 items: 84 in BDI-II, 20 in SDS, 25 in RDI, 34 in PVP, 52 in CRS, 31 in SAD, 20 in CES-D.

The attributes of the clinical context were obtained from the DSM-5 (15), Beck's theory (2), Seligman's theory (1), and the literature (2) for a total of 20 attributes, which are placed in the columns of the matrix.

In this way it is possible to find out what attributes belong to each item, what attributes describe any particular object, and to identify relationships of great clinical and formal importance among objects and attributes (Serra et al., 2015a). Two experts in the field of depression built the clinical context (i.e., the Boolean matrix). The two Psychologists were asked to fill independently a Boolean matrix with the items in rows and the attributes in columns. Whenever an item, in their opinion, investigated a specific attribute (symptom), the corresponding cell in the matrix should have been filled with 1, otherwise with 0. The agreement was calculated using Cohen's K coefficient, and the value was 0.83 indicating a good agreement between experts. The remaining disagreements were discussed and solved by means of a focus group between the experts.

Applying the FPA, four different configurations that may occur within the matrix deserve a separate description, since they produced important modifications in the number of items to be included in the final model:

1. Items that investigate none of the attributes for depression are not useful for the measurement of the construct (their row in the matrix will contain only zeroes). In this case the items were not considered sufficiently precise for the depression construct by the two experts. However, if an item investigated an important attribute that was not included in Table 4.1, it was taken into account by the experts in the final evaluation.

2. Different items investigating the same set of attributes form <u>equivalence classes</u>. It is then useful to choose the items that relate better with the investigated attributes.

3. Attributes not investigated by any item necessitate the construction of new ad hoc items to investigate them. Attributes (i.e. symptoms) for MDE were chosen by experts in the field of mood disorders, so the selected clinical criteria were all considered essential for evaluating the construct.

4. Some items present problems with phrasing, construction, or validity. In this case, the experts could not consider that specific item by explaining the reason.

The resulting formal context is the starting point for the evaluation of a single item and of a single self-report questionnaire. Moreover the clinical context is the reference point for the future construction of a new tool through FPA (the description of this following passage will be the core of the next Chapter which is the second research of this work).

3. Results

3.1 The clinical context

The clinical context is the first result. Therefore, from Table 4.2, it can be seen that the first key result is that we were able to get 30 equivalent classes collecting the same information redundantly investigated by the initial 266 items.

Conditions 1, 2, and 4, described in the previous section, allowed:

✓ To group many questions repeated with different words, in the various self-report questionnaires consulted. Items that investigated the same attributes, and consequently formed equal rows, were replaced by a single item that included the set of attributes belonging to the items. The item was chosen as the most representative of the equivalence class.

✓ To eliminate some items since they did not investigate any of the selected attributes; they were not considered part of the depression construct.

✓ To exclude others items because of problems with their phrasing (double sentences, fuzzy adverbs). The two professionals have judged some items as potentially difficult to interpret for patients and therefore potentially unsuitable for evaluation.

Through these procedures, we were able to get a much more malleable matrix with 30 equivalent classes covering all the identified diagnostic criteria.

Another interesting result of the application of FPA to the set of items was the identification of several items with methodological problems such as:

- Double phrases ("I'm depressed" or "I often want to cry") with the consequent problem of investigating separately in the same item different attributes making the patient's response questionable.

- Fuzzy adverbs ("my life is pretty full").

- On the other hand, problems with content validity (item CRS.40: "I got sick because of the bad weather we have been having").

Another important key finding is that the matrix allowed analysis of the equivalent classes of items and their attributes; some classes investigate subsets of attributes assessed by others. In this way, a prerequisite relationship among different classes is derived. For instance, RDI-10, "I feel quite useless," which investigates feelings of worthlessness, is a subset of RDI.2, "I feel a burden to others," which also contains feelings of guilt; and these two items are prerequisites for SDS.19, "I feel that others would be better off if I were dead," which contains feelings of worthlessness and guilt, and thoughts of death.

The relationships created among items in the matrix generate the clinical structure.

Many other inclusion relations among equivalent classes were observed and can be derived from Table 4.2. All these relations are critical because they describe FPA's adaptive reasoning, suggesting the possibility of applying prerequisite relations in

the clinical context. In fact, if the response to the prerequisite of a particular item is negative, a positive response to the other item will be logically excluded (as explained in the previous Chapter).

Table 4.2 displays the clinical context containing the thirty items and the twenty attributes. It has to be stressed that this representation, that is convenient for explanatory purpose, is equivalent to the Boolean matrix with thirty rows and twenty columns explained in the FPA section and in this research.

Table 4.2: The clinical context containing the thirty items and the twenty attributes.

ID	Item	Text	Attributes
I1	BDI-II.24	I feel like I am being punished	A 11
I2	BDI-II.42	I feel more restless or wound up than usual	A7
I3	BDI-II.47	I have lost most of the interest in other people or things	A2, 20
I4	BDI-II.51	I have much greater difficulty in making decisions than I used to	A13
I5	BDI-II.58	I have less energy than I used to have	A9
I6	BDI-II.63	I sleep a lot more than usual	A6, 8, 9
I7	BDI-II.74	I can't concentrate as well as usual	A12
I8	BDI-II.82	I am less interested in sex than I used to be	A2, 3
I9	ZUNG.1	I feel down-hearted and blue	A1
I10	ZUNG.3	I have crying spells or feel like it	A1, 7
I11	ZUNG.4	I have trouble sleeping at night	A6
I12	ZUNG.15	I am more irritable than usual	A 19
I13	ZUNG.19	I feel that others would be better off if I were dead	A14, 10 11
I14	RDI.4	I do not really want to eat	A5
I15	RDI.7	I do not want to do anything	A2 9
I16	RDI.8	I seem to have lost interest in the future	A2, 17, 20
I17	RDI.10	I feel quite useless	A10
I18	RDI.23	I feel a burden to others	A10, 11
I19	PVP.23	the speed of my thinking seems to be reduced	A8, 12
I20	PVP.28	I think of the families and friends who have died	A14
I21	PVP.31	I have made a suicide attempt	A15
I22	CRS.2	I am losing weight	A4
I23	CRS.12	Dying is the best solution for me	A14, 15
I24	CRS.19	I wake up often in the middle of the night	A6, 7
I25	CRS.21	I am so slowed down that I need help with bathing and dressing	A8, 9, 20
I26	CRS.47	I get hardly anything done lately	A 8, 9, 18
I27	CRS.48	There is only misery in the future for me	A 17
I28	SAD.30	I have the impression of being aloof and not to feel affection for my family members	A 20
I29	CES-D.3	I felt that I could not shake off the blues, even with help from my family or friends	A1, 18
I30	CES-D.15	People were unfriendly	A16

An important observation concerns RDI.2, "I feel better in the evening than in the morning." It is difficult to assign attributes to this item, but despite this, it is very representative of depressive symptoms. Perhaps it would be appropriate to create an ad hoc attribute for this item.

Thanks to the analysis of all the items of the seven self-report tools, we achieved different aims: The clinical context obtained by the 30 chosen items of the total 266, which investigate all the clinical criteria selected, represents the starting point for the construction of a new tool. Indeed, these items contain all the clinical information considered by us to be important for the assessment of depression; moreover, we have also been able to observe potential new attributes (symptoms) to consider in the next step (the construction of a new tool trough FPA methodology). Finally, the first clinical context obtained by the total 266 items allowed us to underline the weaknesses of the seven self-report tools.

3.2 Self-report questionnaires analysis.

None of the explored questionnaires could cover all the attributes for depression alone.

- ➢ BDI-II does not provide information concerning change in weight, Beck's negative view of the world and Seligman' learned helplessness.
- ➢ In SDS, there are no items investigating psychomotor retardation, possible feelings of guilt, possible suicidal attempts, or thoughts of death and Seligman' learned helplessness. Some symptoms are part of the diagnostic criteria of the DSM-5 and, despite their obvious importance, are not considered.
- ➢ Even in RDI, some of the attributes derived from the DSM-5 are not investigated: weight modifications, the decrease in sexual interest and pleasure,

psychomotor retardation, thought of death, possible suicidal attempts and Seligman' learned helplessness.

➢ PVP was built to create an ad hoc self-report questionnaire for depression investigating all DSM diagnostic criteria, however, it does not take into account negative view of the world, negative expectation for the future of beck theory, and Seligman' learned helplessness.

➢ The only attribute missing in CRS is indecision. Nevertheless, some items have problem of content validity (i.e. item CRS.40: "I got sick because of the bad weather we have been having").

➢ SAD does not take into account psychomotor retardation and suicidal ideation or attempts.

➢ Finally, the CES-D does not investigate decreased interest in sex, change in weight, indecision, recurrent thoughts of death, suicidal ideation or attempts and irritability

4. Discussion

This work has shown how the FPA highlights each self-report questionnaire's strengths and weaknesses in terms of correspondence to a set of diagnostic and clinical criteria. The FPA details the relations between objects (items) and attributes (decomposition of clinical and diagnostic criteria). This methodology allows to eliminate useless redundancy and to increase efficiency. FPA also allows for the pinpointing of the relations among sets of items and attributes by analyzing the presence or absence of diagnostic criteria in the items.

Flexibility is another crucial advantage of FPA: the set of attributes could be easily modified or updated according to new versions of DSM or to different theoretical approaches, while the methodology remains equally effective and reliable.

One important result of this first research is the identification of 30 equivalent classes representing the basis for an assessment tool for depression. Such an instrument would explore all the selected diagnostic criteria in term of attributes, without redundancy, and would provide the clinician with a clear reference between items and construct criteria to mimic the interview procedure. The careful analysis of the items of the questionnaires has allowed creating the skeleton for the construction of a new instrument for the major depressive episode.

The strong innovation of FPA comes from the construction of the matrix that allows for the identifying of the actually existing relations among items in terms of the clinical symptoms they endorse. As stated before, such information is already present in the items, but it is hidden by a classical testing methodology that considers the score the most relevant output that questionnaire provides.

The matrix can be expressed in terms of the clinical structure that is the core of the methodology. The structure is the set representation of the implications among the items of the domain. Indeed, it contains all the clinical states (see Chapter 3). The prerequisite relation allows for adaptivity, just as, in a semi-structured interview, the individual is driven to respond to items according to what he answered previously. For example, in the case of depression, if a patient answers "no" to an item relating to "thoughts of death," the adaptive algorithm of FPA will not investigate whether he intends to die by suicide because "having thoughts of death" is a prerequisite of the suicide attempt. In this way, the tool becomes adaptive because it allows for a thorough analysis of the areas in which the patient suffers.

This work showed the multiple advantages of FPA methodology. It avoids redundancy and unnecessary collection of information saving time and energy. Moreover, the clinician obtains qualitative information about a patient's symptoms in a systematic and methodologically solid framework (Serra et al., 2015a). Indeed, different response patterns (i.e., different attribute configurations) may characterize people who obtain the same scores on a self-report questionnaire. The information can be used by FPA to detect differences among these people, and produce specific indicators that could be used when planning treatments (Bottesi, Spoto, Freeston, Sanavio, & Vidotto, 2015a). Specific psychological mechanisms underlying each patient's phenomenology are thought to have implications for treatment effectiveness (Serra et al., 2015a).

Different combinations of symptoms could produce the same score on a self-report questionnaire, although such information might not be regarded in clinical practice. Indeed, considering two individuals who obtained similar scores on the Somatic-Affective Scale of BDI-II, such scores may arise predominantly from an elevation in either somatic or affective features. BDI-II does not allow for discrimination between the two cases. On the contrary, FPA is useful in clarifying the specific clinical configuration depicted by the observed response pattern, rather than the mere score.

Summarizing, the FPA, through its methodology, allows for the construction of new clinical tools for clinical evaluation following efficient and effective principles beyond the assessment of depression. In this particular case, starting from several self-report questionnaires and numerous diagnostic criteria considered essential for the assessment of depression, the FPA applied to depression' questionnaires allowed for the creation of a start point to the construction of a new tool with many added benefits compared to the self-report questionnaires used in the research.

You can find a short version of this work in "Serra, Spoto, Ghisi & Vidotto, 2015, Plos One".

In the next chapter, the steps for the construction and the validation of a new tool will be discussed. In particular, the new tool built up for the assessment of Major depressive Episode shows the opportunity to fruitfully use the qualitative information already present in the questionnaire, but hidden by the score, which is crucial when it comes to suggesting the elective treatment strategies. Therefore, FPA could represent an important approach for improving case conceptualization and treatment implementation.

CHAPTER 5

The construction and validation of a new tool through Formal Psychological Assessment: *the clinical state* as effective output.

1. *Introduction and research aim*

Over the years, many improvements have been achieved in the various stages of the assessment. In Chapter 2, we reviewed the history of the assessment, highlighting the attempts of improvement and the weaknesses that still exist. In this Chapter, we focus on the limitations found in depression assessment tools, and on the other hand, on the general limits of self-report tools, with the most important goal of finding a solution. Specifically, the tools for evaluating mood disorders, particularly those involving major depressive episodes (MDEs) ones, show some application limits. Some studies highlighted critical issues regarding self-report depression tools and this is crucial for assessment and treatment (Baldessarini, Vieta, Calabrese, Tohen, & Bowden, 2010; Hyman, 2014). As reported in the previous Chapter, Pettersson and colleagues (2015) identified sensitivity and specificity issues. Balsamo and Saggino (2007) found limits in the psychometric properties of six regularly used self-report questionnaires (e.g. Overestimation of symptoms by patients compared to the results obtained in the interviews); moreover, each scale reflected the author's theories, each measuring different aspects of the same construct. This last issue was discussed in the previous Chapter, as it was found in the analysis of seven habitually used self-report questionnaires through the FPA approach. In fact, the analysis of relationship between items and diagnostic criteria highlighted as no questionnaire alone could cover all diagnostic criteria for the evaluation of the Major Depressive Episode.

More in generally, many authors underlined the weaknesses of self-report tools and in particular, the psychometric approach issues.

Gibbons, Clark, Cavanaugh, & Davis (1985) emphasized that the traditional method of scoring can be "wrong" because it is based on assumptions that may be false: it gives equal "weight" to each item, assuming that each item or symptom of a clinical scale represents an equal level of psychiatric severity.

Moreover, as many authors showed, self-evaluating questionnaires allow for a systematic and quick collection of a large amount of information and the avoidance of patient embarrassment. However, they redundantly (non-adaptively) investigate constructs and provide only a quantitative numeric score that does not systematically account for qualitative information (Bottesi et al., 2015a; Fava, Ruini, & Rafanelli, 2004; Shapiro, 1951; Spoto et al., 2013; Wright and Feinstein, 1992).

Based on the assumptions of the studies described above, it is important to consider all of the limits that self-report tools have, by taking into account the possible overestimation of symptoms by patients (Faravelli, Albanesi, & Poli 1986), and the inability of MDE self-report measures to enclose the whole set of depressive symptoms, whether agitated or inhibited (Koukopoulos and Koukopoulos, 1999; Serra et al., 2015a), in the construction of the item. Finally, it is also relevant to remember not to take into account only the patient's cut-off scores (Bottesi et al., 2015a; Fava et al., 2004; Gibbons et al., 1985). As a consequence of the last statement, even overtaking or not overtaking the cut-off may not always be so important. In fact, the cut-off provides only a quantitative score, but if two patients have the same score, it does not mean that they have equal symptomatology (one could be much more serious than the other, since he/she responded positively to more severe symptoms).

The purpose of this study is to create a new tool for Major Depressive Episode (MDE) evaluation to overcome the difficulties of the MDE tools described above. The present

study aims to construct an adaptive-qualitative tool that investigates all MDE clinical features and provides qualitative information (and not just a score) to differentiate patients with the same score but different symptoms as well as differing severities of psychopathology. In this work, we applied again the FPA framework and fruitfully use the main concepts described in chapter 3.

2. Materials and Methods

2.1. Tool Construction

Three important steps were achieved in constructing the new tool:

First, various features and symptoms (clinical criteria) of MDEs were analysed and categorized as attributes of the clinical context. As described in the DSM-5, there are different types of Major Depressive Episodes (MDEs); in particular, a MDE may be part of major depressive disorder (MDD) or bipolar disorder (BD, Type I or II), with symptoms more agitated or more inhibited.

Second, the items (objects of the clinical context) were constructed on the basis of one or more chosen clinical criteria (attributes).

Third, In line with the FPA methodology, the matrix was obtained to analyse all of the relationships among items (objects) and diagnostic criteria (attributes). More specifically, the items of the tool were verified as covering the entire set of clinical criteria (all columns contained at least one "1"). This result was achieved through the agreement of four specialists in the field of mood disorders selected on the basis of their expertise in the field of Cognitive Behavioural Therapy and psychological assessment. More specifically, experts were asked to fill independently a Boolean matrix with the items in rows and the attributes in columns. Whenever an item, in their opinion, investigated a specific attribute, the corresponding cell in the matrix should have been filled with 1, otherwise with 0. The Cohen's k coefficient was computed for each pair of experts' matrices and resulted in an average value of 0.88 indicating a very good

agreement among experts. The remaining disagreements were discussed and solved by means of a focus group.

First of all, MDE is described by the decomposition of the DSM-5 diagnostic criteria (2013) for this disorder (it may be part of both MDD and BD). All of the described diagnostic criteria were taken into account:

- A1 (*depressed mood*), A2 (*diminished interest and pleasure*), A3 (*decreased interest in sex*), A4 (*increase or loss of weight*), A5 (*gain or loss of appetite*), A6 (*insomnia or hypersomnia*), A7 (*agitation*), A8 (*psychomotor retardation*), A9 (*fatigue or energy loss*), A10 (*feelings of worthlessness*), A11 (*feelings of guilt*), A12 (*diminished ability to think and concentrate*), A13 (*indecision*), A14 (*recurrent thoughts of death*), and A15 (*suicidal ideation or attempted suicide*). Criterion A15 underlines the seriousness of "suicidal ideation" in MDE. We decided to separate this symptom by thoughts of death. According to several authors, suicide is indeed the third-highest cause of death in the population between 15 and 35 years old (Baldessarini et al., 2006a; Gunnell and Middleton, 2003) and it is often associated with misdiagnosis. Moreover, suicide is associated with a mood disorder in 90% of cases (Baldessarini et al., 2006a).

- Attributes A16, A17, and A18 are related to two cognitive behavioural theories, Beck's hopelessness theory and Seligman's helplessness theory, which are described in the first Chapter.

After a careful literature review, other clinical criteria for MDE were taken into account because they are widely described in the literature, and they are potentially able to discriminate between different types of MDEs:

- Criterion A19 refers to *irritability* (Henderson et al., 2013; Pedrelli et al., 2013); a person with depression can easily feel frustrated, which often results in anger, crying, nervousness, and mood lability (Benazzi & Akiskal, 2005).

- Criterion A20 refers to *apathy* (Mulin et al., 2011; Alexopoulos et al., 2013). Patients with depression often are characterized by decreased emotional reactions to situations and events in everyday life. Apathy is expressed in the form of indifference, physical inertia, or lack of reaction when facing situations that would normally arouse interest or emotion, as well as a reduction of purposeful behaviour, a lack of initiative, and submission in one's daily choices.

The clinical criteria described above are the same as used in the previous research. In keeping with the previous research and the self-report questionnaires content, (see Chapter 4) we have added three other clinical criteria:

- Criterion A21 refers to *health concerns* (House, 1989; Magariños, Zafar, Nissenson, & Blanco, 2002). It can take on the characteristics of real hypochondria in MDE, and the concerns may be related to somatization disorders.

- Criterion A22 refers to *somatic disorders* (Goodwin and Jamison, 2007; Al Busaidi, 2010; Campo, 2012). It can be expressed through a myriad of symptoms in people with MDE: neuro-vegetative disorders, stomach cramps, vomiting, difficulty of digestion, diarrhoea, palpitations, hyperventilation, paraesthesia, sweating, flushing, tremors, headaches, increased heart rate, an urgent need to urinate often, a feeling of heaviness in the limbs or in the head, and back or muscle pain.

- Criterion A23 was inserted as the last in the list of clinical criteria for assessing MDE. The literature review and the presence of this attribute in the items of

almost all MDE scales analyzed by Serra et al. (2015) demonstrate its important contribution to depression evaluation. In fact, individuals with depression usually feel better at the end of the day when they can go to sleep and do not have to face their daily problems anymore.

Table 5.1 summarizes all of the sets of clinical criteria that we considered in constructing the MDE assessment tool.

Table 5.1. The twenty-three clinical criteria for major depressive episode construct.

Attribute	Explanation
A1	Depressed mood
A2	Diminished interest and pleasure
A3	Decreased interest in sex
A4	Increase or loss of weight
A5	Gain or loss of appetite
A6	Insomnia or hypersomnia
A7	Agitation
A8	Psychomotor retardation
A9	Fatigue or energy loss
A10	Feelings of worthlessness (or Beck's negative view of self)
A11	Feelings of guilt
A12	Diminished ability to think and concentrate
A13	Indecision
A14	Recurrent thoughts of death
A15	Suicidal ideation or attempted suicide
A16	Beck's negative view of the world
A17	Beck's negative expectation of the future
A18	Seligman's learned helplessness
A19	Irritability
A20	Apathy
A21	Health concern
A22	Somatic disorders
A23	More positive mood in the evening

On the base of the criteria described above, 41 items have been constructed. Some of them contain a single diagnostic criterion; for example, the item "I feel helpless in the face of life events" contains A18 (the learned helplessness of Seligman). Other items were constructed to include two or more diagnostic criteria; for example, the item "I

feel nervous about this sadness I never abandon" contains three diagnostic criteria: A17 (Beck's negative expectation of the future), A1 (depressed mood), and A7 (agitation). In line with the FPA methodology, the matrix allowed analysing all of the relationships among these items and diagnostic criteria. The whole set of clinical criteria (attributes) in Table 5.1 were investigated by the set of 41 items. The new perspective of this tool is more deeply explored in the Results and Discussion sections.

2.2.Participants

The research participants who were tested were divided into clinical and non-clinical groups.

The clinical group consisted of 38 subjects with MDE (who were diagnosed with major depressive disorder or bipolar disorder, or else with MDE during their first access in the day hospital) of the Neurosciences, Mental Health, and Sensory Organ (NESMOS) Department of La Sapienza University, Rome. In particular, the patients included in the study comprised eight individuals who were on their first access to the day hospital, four people who had a reactive MDE (caused by a stressful event or a death event), two who had an MDE with familiar genetics in mood disorders highlighted by their medical history, and two who suffered from unspecified MDE. Eleven patients were suffering from MDE within a major depressive disorder; one patient of this group had comorbid obsessive compulsive disorder (OCD), one patient had comorbid social anxiety disorder, and two others had comorbid eating disorder (anorexia nervosa). Nine patients were suffering from an MDE within bipolar disorder type I; two of them had comorbid OCD, and two other patients had comorbid eating disorder (anorexia nervosa and bulimia). Finally, ten patients were suffering from MDE in bipolar disorder type II; three of them had comorbid OCD, two of them had comorbid panic disorder, and one of them had comorbid social anxiety disorder. The exclusion criteria were mental

retardation and psychotic traits to avoid problems in interpreting the responses to the QuEDS. Of the participants, 47% were male and the remaining 53% were female. A majority of the participants had a high school diploma, and their ages ranged between 21 and 69 years.

The non-clinical group consisted of 265 Italian individuals from different regions. The convenience non-clinical sample of the present research included individuals recruited in the area of the University of Padova (both students and non-students). The exclusion criteria in the non-clinical group involved all individuals suffering from MDE (e.g., those who were under pharmacological or psychotherapeutic treatment for depression). Among these participants, 70% were female. A majority of participants had a high school diploma, and their ages ranged between 19 and 56 years.

2.3. Clinical Tools

- ➢ ***The Qualitative–Quantitative Evaluation of Depressive Symptomatology (QuEDS)***. The QuEDS tool (Serra et al., 2017) contains 41 dichotomous items constructed on the basis of 23 clinical criteria of major depressive episodes from the DSM-5 and the literature. The maximum score is 41, and the minimum is 0. It is assumed that if a person responds positively to an item, then he/she has the symptoms (in terms of clinical criteria or attributes) included in this item. The respondents are asked to reply "yes" or "no" to indicate the presence of symptoms.

- ➢ ***Depression-Anxiety-Stress-Scale 21 (DASS-21)***. The DASS-21 is the short version of the self-report test designed to measure the three related negative emotional states of depression, anxiety, and stress (Bottesi, Ghisi, Altoè, Conforti, Melli, & Sica 2015b; Henry and Crawford, 2005). DASS-21 contains seven items for assessing depression, seven items for assessing anxiety and

seven items for assessing stress. The Depression scale evaluates dysphoria, hopelessness, devaluation of life, self-deprecation, lack of interest/involvement, anhedonia, and inertia. The Anxiety scale assesses autonomic arousal, situational anxiety, and subjective experience of anxious affect. The Stress scale (items) is sensitive to levels of chronic non-specific arousal. It assesses difficulty in relaxing as well as being easily agitated, irritable, and impatient. The respondents are asked to use a 4-point Likert scale to indicate the severity and frequency of symptoms.

2.4. Procedure and Administration

All of the research participants completed informed consent and sociodemographic forms before answering the questionnaire items. No time limit was imposed to complete the questionnaires. All 265 subjects of the non-clinical group completed the QuEDS for major depressive episodes.

A subgroup of 113 individuals of the non-clinical group also answered the self-report measure DASS-21 to evaluate the convergent and divergent validity of the QuEDS. Moreover, 63 out of these 113 subjects compiled the QuEDS twice, after 1 month, to evaluate the temporal stability of the scale (test–retest).

The NESMOS Department's psychiatrists, diagnosed the patients of clinical group as Major Depressive Episode patients through a depression rating scale (SCID-I). Then these patients responded to the QuEDS.

At clinic intake, participants provided written, informed consent for potential research analysis and anonymous reporting of clinical findings in aggregate form, in accord with Italian legal and ethical requirements. The study was conducted in accordance with the Declaration of Helsinki. All participants entered the study of their own free will and they were informed in detail about the aims of the study, the voluntary nature of their

participation, and their right to withdraw from the study at any time. Furthermore, participants were allowed to ask for restitution about their own score, providing authors with their own auto generated code, used during the administration phase.

2.5. Data Analysis

The data of all 303 participants were used for analysis, no missing data were observed. Different kinds of data analyses were conducted to test the validity and reliability of the new tool (QuEDS).

Inferential analyses were conducted by means of the software R 3.3.0 (R Core Team, 2013), while confirmatory factor analysis (CFA) was conducted by means of the software LISREL 8.80 (Jöreskog, & Sörbom, 1986; 1989; 1993). The discriminant validity was evaluated by means of both a comparison between the scores of non-clinical and clinical groups, and referring to the classical ROC curves approach. Convergent validity was evaluated by computing the correlation between the QuEDS and the DASS-21 scores. The reliability of the scale was tested both with respect to the internal consistency and to the test-retest. Content validity has been evaluated by referring to the FPA methodology. Finally, again by means of the FPA, the capability of the tool to clinically discriminate patients has been tested and reported.

3. Results

3.1.Construct Validity

The construct validity of the QuEDS was evaluated by investigating its factorial validity, discriminant validity, and convergent–divergent validity.

> *Factorial validity*: From a conceptual point of view, depression is a strong and united construct; some authors, including Beck, attributed more importance to the cognitive dimension of depression (Rainone & Mancini, 2007) without neglecting the somatic-affective dimension. Many authors identified three

dimensions in the depression construct: cognitive, somatic, and affective (Dinger et al., 2015; O'Hara, Neunaber, & Zekoski, 1984; Roca, Wigley, & White, 1996). A hierarchical model having three sub-dimensions (namely: cognitive, affective and somatic) and one second order factor (depression) could be the factorial solution that is most likely to represent the structure of the investigated construct; in fact, this model would have the advantage of offering an interesting explanation, from the clinical perspective. Indeed, the QuEDS was constructed both specifically for evaluating depressive symptoms and sensitively to evaluate thoughts (cognitive), somatic aspects (somatic), and emotions (affective) related to depression. The authors are interested in testing, on the non-clinical sample (Osborne & Costello, 2004), a hierarchical factorial structure (Berrios, Kellett, Fiorani, & Poggioli, 2015; Roberts, Hart, & Eastwood, 2015) with three sub-factors (i.e., cognitive, affective, and somatic), all linked to a second-order factor (i.e., depression). The clinical sample for factor analysis was not considered due to the low number of participants (i.e. N = 38). Furthermore, in order to compare the fit of this model to different theoretically plausible solutions, it was compared with three other different factorial models to the collected data:

1. A model with one latent construct which we called "depression".

2. A model with two latent factors, which we called "somatic-affective factor" and "cognitive factor".

3. A model with three factors: cognitive, somatic and affective factors.

   In the proposed hierarchical model, the items were grouped into the sub-factors as follows:

- *The cognitive factor* includes items 5, 6, 9, 10, 14, 19, 20, 21, 24, 25, 27, 30, 32, 33, and 41. It comprises symptoms related to distortions of thought systems and also to feelings of guilt, helplessness, worthlessness, hopelessness, and death.

- *The somatic factor* includes items 1, 2, 3, 4, 11, 13, 16, 22, 23, 26, 28, 31, 35, and 39. It comprises symptoms related to fatigue, sleep, appetite, psycho-motor retardation, and other somatic disorders that often involve MDEs.

- *The affective factor* includes items 7, 8, 12, 15, 17, 18, 29, 34, 36, 37, 38, and 40. It comprises the emotions that characterize different types of MDEs, including sadness, apathy, irritability, agitation, and various concerns.

Table 5.2 displays a comparison of the fit indexes for the four tested factorial structures. The table shows that the hierarchical model fits the data better than any of the three other models. All of the fit indexes for the hierarchical model (with no use of modification indexes) had adequate values. More specifically, the ratio between the Chi-square and the degrees of freedom, the RMSEA, the CFI, and the NNFI showed a good fit, while the NFI indicated an adequate model fit (Marsh, Hau, & Wen 2004). Furthermore, no significant double loadings were observed, nor correlation among error terms.

Table 5.2. The fit indexes of the three tested models.

Fit Index	Mono-factorial	Two-Factor Model	Three-Factor Model	Hierarchical Model
χ^2/df	2.67	2.51	2.54	2.45
RMSEA	0.07	0.07	0.07	0.06
NFI	0.93	0.94	0.94	0.94
NNFI	0.95	0.96	0.96	0.96
CFI	0.96	0.96	0.96	0.96
SRMR	0.08	0.08	0.08	0.07
AIC	2204	2121	2141	2068

For the hierarchical model, all of the items' saturations on the respective factors were significant and ranged between 0.26 and 0.76 for the cognitive factor; between 0.32 and 0.71 for the somatic factor; and between 0.33 and 0.70 for the affective factor (Table 5.3). The results support the selection of the hierarchical model by confirming that its underlying factorial structure has a higher-order factor accounting for the relationship among lower-order specific factors (Subica et al., 2014). In second-order models, it is

necessary for the lower-order specific factors to be correlated among each other and with the higher-order factor (Schmid and Leiman, 1957). In this specific case, the links between the sub-factors and the higher-order factor were in the range of 0.70–0.91, once more supporting the selected model.

Table 5.3. Factor loadings of each of the 41 items of the QuEDS. No double loading were observed. In the last line of the table are displayed the strengths of the links between the first and second-order factors.

	Cognitive Factor	Somatic Factor	Affective Factor
Item 5	.68	-	-
Item 6	.53	-	-
Item 9	.54	-	-
Item 10	.65	-	-
Item 14	.76	-	-
Item 19	.56	-	-
Item 20	.75	-	-
Item 21	.53	-	-
Item 24	.78	-	-
Item 25	.80	-	-
Item 27	.45	-	-
Item 30	.66	-	-
Item 32	.72	-	-
Item 33	.45	-	-
Item 41	.82	-	-
Item 1	-	.61	-
Item 2	-	.47	-
Item 3	-	.71	-
Item 4	-	.44	-
Item 11	-	.33	-
Item 13	-	.58	-
Item 16	-	.61	-
Item 22	-	.53	-
Item 23	-	.51	-
Item 26	-	.59	-
Item 28	-	.76	-
Item 31	-	.57	-
Item 35	-	.68	-
Item 39	-	.44	-
Item 7	-	-	.41
Item 8	-	-	.30
Item 12	-	-	.77
Item 15	-	-	.74
Item 17	-	-	.50
Item 18	-	-	.80
Item 29	-	-	.64
Item 34	-	-	.72
Item 36	-	-	.70
Item 37	-	-	.69
Item 38	-	-	.53
Item 40	-	-	.39
Depression	.91	.87	.92

➢ *Discriminant validity*: The scores of 38 patients with MDE were compared with those of the 265 non-clinical subjects to test the discriminant validity of the

QuEDS. The clinical group obtained an average score of 28.5 (sd = 6.5), while the non-clinical group obtained an average score of 6.5 (sd = 6). The difference between the two groups, tested using a t-test for independent samples, was significant (t_{299} = −20.20; p < .001) and supported the validity of the QuEDS. Furthermore, in order to test the ability of the QuEDS in separating the two groups, an analysis based on the ROC curves has been carried out. Results showed a very good value of the AUC statistic (confidence interval 0.97 - 0.99). Moreover an optimal threshold score of 19 was determined that allowed for a specificity of .98 and a sensitivity of .94.

- ➤ *Convergent–divergent validity.* The convergent validity of the QuEDS was verified by comparing its scores in the non-clinical sample with those of the DASS-21 (which, as described above, is constituted by three sub-scores for Depression, Anxiety, and Stress). We chose to use the DASS-21 as considered the most suitable tool for the specific case thanks to the rapidity of administration and sensitivity in the measurement of all the three constructs (Depression, Anxiety and Stress). The correlations among the 113 subjects' scores in the QuEDS and the three sub-scales of DASS-21 were all significant. More specifically, the correlation between the QuEDS and the Depression subscale of the DASS-21 was r = .72 (p < .05); the correlation between the QuEDS and the Anxiety subscale of the DASS-21 was r = .39 (p < .05); and the correlation between the QuEDS and the Stress subscale of the DASS-21 was r = .59 (p < .05). These results are not surprising, since the depression construct may have several features in common with the stress and anxiety constructs (Bayram, & Bilgel, 2008). However, it has to be stressed that the correlation between the QuEDS and the Depression subscale was significantly higher than the

correlation between the QuEDS and the Anxiety subscale (z = 3.32, p < .001); on the contrary, a not significant difference was observed between the correlation of the QuEDS with the Depression subscale and the correlation of the QuEDS and the Stress subscale (z = 1.71, n.s.). These results indicate a good divergent validity of the QuEDS. The correlations among the sub-factors (cognitive-somatic-affective) of the QuEDS and the subscales of the DASS-21 have also been computed, and the results are displayed in Table 5.4. While the correlation between the Anxiety subscale of the DASS-21 and the factors of the QuEDS was systematically and significantly lower than the correlations between the QuEDS subscales and the Depression subscale of the DASS-21 (cognitive: z = 2.57, p < .01; somatic: z = 2.24, p < .05; affective: z = 2.03, p < .05), the situation was the opposite with respect to the Stress subscale. In fact, all the correlations between the three sub factors of the QuEDS and the Depression subscale of the DASS-21 were not significantly higher than their correlation with the Stress subscale (cognitive: z = 1.65, n.s., somatic: z = 1.23, n.s.; affective: z = −0.11, n.s.). Table 4 displays the 7x7 correlation matrix of the QuEDS total score, QuEDS subscales, and the subscales from the DASS-21.

Table 5.4. The correlation matrix of the 3 subscales from the DASS-21, the three subscales of the QuEDS and the total score of the QuEDS.

	DASS-21 Depression	DASS-21 Anxiety	DASS-21 Stress	QuEDS-cognitive	QuEDS-somatic	QuEDS-affective	QuEDS-TOT
DASS-21 Depression	-						
DASS-21 Anxiety	0.37	-					
DASS-21 Stress	0.44	0.50	-				
QuEDS-cognitive	0.64	0.37	0.48	-			
QuEDS-somatic	0.53	0.25	0.40	0.37	-		
QuEDS-affective	0.58	0.34	0.56	0.61	0.46	-	
QuEDS-TOT	0.72	0.39	0.59	0.80	0.78	0.84	-

3.2 Reliability

A reliability analysis, on the sample, showed that the QuEDS scale has a very good internal consistency, Cronbach's α = .948. Even the alpha values relative to the three sub-factors were good: cognitive factor (α = .91; average item inter-correlation = .41), somatic factor (α = .86; average item inter-correlation = .31), affective factor (α= .82; average item inter-correlation = .41). Given such values and the number of items in the scale, the alpha precision can be considered adequate (Cortina, 1993).

Regarding the test–retest reliability of the QuEDS, the correlation among the scores of the 63 test subjects at Time 1 and Time 2 (after one month) was .74, which indicates good stability for the tool.

3.3.Content validity

The QuEDS was created to answer the following question: Does it include the most common symptoms related to various types of MDEs? The FPA methodology was used to answer this question. A matrix was created with 41 items in the rows and 23 clinical criteria in the columns, called the "clinical context." Thus, it was verified that each item would include one or more clinical criteria. Four specialists in the field of mood disorders carried out this analysis. Table 5.5 shows the content of the items and the set of attributes (symptoms) that each item investigates.

Concerning content validity, the first key result is that the QuEDS was able to collect all of the information from 41 items, in terms of clinical criteria, to evaluate different types of MDEs. In addition, the items in the Table 5.5 may include one or more clinical criterion. The previous research (Chapter 4) highlights that none of the widely used tests to assess depression alone is able to investigate all clinical criteria, even those related to the DSM-5 (Serra et al., 2015a).

Table 5.5 represents the clinical context in a different way to Boolean matrix.

Table 5.5. The items with their attributes

Items	Clinical Criteria
1. I feel that I don't have the same energy to have sex	A3, A9
2. I often wake up in the middle of the night and I can't asleep again	A6, A7
3. I feel like that my thinking is slowing down	A8, A12
4. I have sleeping problems	A6
5. I am stressed by feeling of guilt	A11
6. I am think the world is cruel and unhappy	A16
7. I keep crying very easily	A1, A7, A19
8. I get irritated very easily	A19
9. I think my life is hell and I only deserve to feel bad	A1, A11, A16
10. I fell incapable to face life's events	A18
11. I suffer of somatic disorders (e.g. headache stomach ache)	A22
12. I have lost interest in the future which doesn't save anything good for me	A2, A17
13. I am less interest in sex	A3, A20
14. I feel incapable and totally useless	A10, A18
15. I see the same unhappiness I have now in the future	A1, A17
16. My desire to eat is not the same	A5
17. I often feel like crying, but I cannot do it	A1, A20
18. I cannot have any interest and pleasure in people and things that before I was interested in	A2, A20
19. I thought to kill my self	A14, A15
20. Sometimes I think it would be better if I were dead	A14
21. I am really worried about my health	A21
22. My weight has had significant changes	A4
23. I've visibly lost (or gained) weight	A4
24. I am afraid of about everything that it will happen to me because I am not able to do anything	A17, A18
25. I feel like I don't have any more power over my empty and sad life	A1, A16, A18
26. My appetite has changed	A5
27. To make choices is hard for me	A12, A13, A20
28. I feel I 'm slowing down in my daily routines	A8, A9
29. I feel helpless and inhibited facing my incapacity to concentrate	A12, A20
30. I feel too much on the other people that it would be better if I killed myself	A10, A11, A14
31. I have not much energy and I feel tired	A9
32. I am disappointed of myself and the choices I made	A10, A11
33. I have problems in making decisions	A13
34. I feel sad	A1
35. My ability to think and memorize has been reduced	A8, A9, A12
36. I don't have any interest and desire in doing anything	A2
37. I am agitated of the idea that this sadness won't ever leave me	A1, A7, A17
38. I feel agitated	A7
39. I feel so tired and without any energy that I need help to wash myself and to get dressed	A8, A9, A18, A20
40. I am better in the evening more than in the morning	A23
41. I often feel like a loser	A10

Furthermore, in formulating the items, we avoided methodological problems such as double phrases ("I'm depressed" or "I often want to cry"), fuzzy adverbs ("my life is pretty full"), or problems with content validity. In other words, each item of the QuEDS includes one or more clinical criteria described in Table 5.1; none of the items investigate other symptoms that may be related to depression but are not part of the construct (e.g., items about anxiety, obsessions). In the matrix, this means that there were no empty rows (with all "0"). Moreover, there were no empty columns in the matrix; this means that the 41 items of the QuEDS investigate all 23 of the clinical criteria in Table 5.1.

Items 22 and 23, as well as items 16 and 26 have the same attribute to check that the reply to the item is valid since these attributes are not present in other items of the tool.

By applying FPA, it became possible to conduct a content analysis even for the three sub-factor included in the model. For each factor, it has been possible to create a clinical context including all of its items and the subset of attributes investigated by the items of the sub-factor. The results of this procedure are displayed in Table 5.6.

Table 5.6. The clinical contexts of QuEDS three factors. Each row is an item, while each column is a clinical criteria either investigated or not by the item. Every time an item investigates a specific criterion, the corresponding cell will contain "1," (otherwise "0)."

<div align="center">

COGNITIVE FACTOR

</div>

	A1	A10	A11	A12	A13	A14	A15	A16	A17	A18	A21
Item 5	0	0	1	0	0	0	0	0	0	0	0
Item 6	0	0	0	0	0	0	0	1	0	0	0
Item 9	1	0	1	0	0	0	0	1	0	0	0
Item 10	0	0	0	0	0	0	0	0	0	1	0
Item 14	0	1	0	0	0	0	0	0	0	1	0
Item 19	0	0	0	0	0	1	1	0	0	0	0
Item 20	0	0	0	0	0	1	0	0	0	0	0
Item 21	0	0	0	0	0	0	0	0	0	0	1
Item 24	0	0	0	0	0	0	0	0	1	1	0
Item 25	1	0	0	0	0	0	0	1	0	1	0
Item 27	0	0	0	1	1	0	0	0	0	0	0
Item 30	0	1	1	0	0	1	0	0	0	0	0
Item 32	0	1	1	0	0	0	0	0	0	0	0
Item 33	0	0	0	0	1	0	0	0	0	0	0
Item 41	0	1	0	0	0	0	0	0	0	0	0

SOMATIC FACTOR

	A3	A4	A5	A6	A7	A8	A9	A12	A18	A20	A22
Item 1	1	0	0	0	0	0	1	0	0	0	0
Item 2	0	0	0	1	1	0	0	0	0	0	0
Item 3	0	0	0	0	0	1	0	1	0	0	0
Item 4	0	0	0	1	0	0	0	0	0	0	0
Item 11	0	0	0	0	0	0	0	0	0	0	1
Item 13	1	0	0	0	0	0	0	0	0	0	0
Item 16	0	0	1	0	0	0	0	0	0	0	0
Item 22	0	1	0	0	0	0	0	0	0	0	0
Item 23	0	1	0	0	0	0	0	0	0	0	0
Item 26	0	0	1	0	0	0	0	0	0	0	0
Item 28	0	0	0	0	0	1	1	0	0	0	0
Item 31	0	0	0	0	0	0	1	0	0	0	0
Item 35	0	0	0	0	0	1	1	1	0	0	0
Item 39	0	0	0	0	0	0	1	0	1	1	0

AFFECTIVE FACTOR

	A1	A2	A7	A12	A17	A19	A20	A23
Item 7	1	0	0	0	0	1	0	0
Item 8	0	0	0	0	0	1	0	0
Item 12	0	1	0	0	1	0	0	0
Item 15	1	0	0	0	1	0	0	0
Item 17	1	0	0	0	0	0	1	0
Item 18	0	1	0	0	0	0	1	0
Item 29	0	0	0	1	0	0	1	0
Item 34	1	0	0	0	0	0	0	0
Item 36	0	1	0	0	0	0	1	0
Item 37	1	0	1	0	0	0	0	0
Item 38	0	0	1	0	0	0	0	0
Item 40	0	0	0	0	0	0	0	1

It is noteworthy how the sets of attributes investigated by each factor are different. More specifically, only the cognitive factor investigates feelings of worthlessness and guilt, indecision, recurrent thoughts of death and suicidal ideation, Beck's negative view of the world, and finally health concerns; in fact, all of these symptoms are related to thoughts. The somatic factor alone includes decreased interest in sex, increased or loss of weight, gain or loss of appetite, insomnia or hypersomnia, psychomotor retardation, fatigue or energy loss, and somatic disorders; all of these manifestations are physical dysfunctions. Instead, the affective factor alone comprises diminished interest and pleasure, irritability, and a more positive mood in the evening. Finally, some symptoms are investigated by two sub-factors, and one of them is investigated by all of the sub-factors.

The following table briefly describes the links between the sub-factors and the attributes (symptoms) that they share.

Table 5.7. The factors that investigate each single attribute of the clinical context.

FACTORS	ATTRIBUTES
All three factors	A12
Cognitive & affective	A1, A17
Cognitive & somatic	A18
Somatic & affective	A7, A20
Cognitive	A10, A11, A13, A14, A15, A16, A21
Somatic	A3, A4, A5, A6, A8, A9, A22
Affective	A2, A19, A23

The analysis of relationships between items and symptoms allows clinicians to go beyond the score, acquiring qualitative information on the individual, and understanding the patient's symptomatic areas. A descriptive example about how the FPA could integrate the quantitative information collected through the questionnaire is presented below.

3.4. Beyond the Numeric Score: The "Clinical State" of the Patient

The new QuEDS allows clinicians to go beyond the numeric score and focus their analysis on the symptoms that patients experience or about which they complain (Serra et al., 2017).

It may be useful to introduce a practical example from the patients of this study; two of the 38 patients in the clinical group were chosen. They obtained the same score on the QuEDS: 31. This means that both answered "yes" to 31 items out of the 41 total items. This score is clearly high, and the patients, who had already been diagnosed with MDE, were confirmed to have a severe depressive symptomatology with this scoring. However, the two patients did not have exactly the same disorder. Patient SC was suffering from a reactive MDE (subsequent to a stressful life event), while patient FG

was suffering from MDE inside bipolar disorder type 1 (see chapter one). As stated in first chapter, many authors (e.g., Benazzi & Akiskal, 2005; Koukopoulos and Koukopoulos, 1999; Maj et al., 2003; etc.) reported that MDEs in bipolar disorder often occur with mixed or agitated features. According to the classical methodology, the questionnaire's output is the same for both patients. In agreement with several other authors (Bottesi et al., 2015a ; Fava et al., 2004; Gibbons et al., 1985; Serra et al., 2015a Wright and Feinstein, 1982), in this study, it was assumed that if the two patients had the same score, this did not mean that they had equal symptomatology. Indeed, they may have answered affirmatively to the same number of items but not to the same items, and the whole symptomatology may be more serious in one of them. Unlike in the usual methodology, qualitative information on the two patients' symptoms were collected through their clinical state.

Patient SC responded affirmatively to items 1, 3, 5, 6, 10, 12, 13, 14, 15, 16, 17, 18, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 39, 40, and 41. Consequently, his clinical state contained the following symptoms (attributes) in terms of clinical criteria: A1, A2, A3, A4, A5, A8, A9, A10, A11, A12, A13, A14, A16, A17, A18, A20, and A23. Patient FG responded affirmatively to items 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 24, 25, 26, 27, 29, 32, 33, 34, 36, 37, 38, 40, and 41. Accordingly his clinical state contained the following symptoms (attributes) in terms of clinical criteria: A1, A2, A3, A5, A6, A7, A10, A11, A12, A13, A14, A15, A16, A17, A18, A19, A20, A21, A22, A23 (to see which criteria the items are related to, see Tables 5.1 and 5.5).

The listed attributes of both patients comprise the items they answered positively. As can be seen, the two patients had two different clinical states. Specifically, they shared a large number of attributes (namely, A1, A2, A3, A5, A10, A11, A12, A13, A14, A16, A17, A18, A20, and A23). This fact indicated that many of the general characteristics of

the affective episode of both patients were the same. Nevertheless, each patient presented some specific characteristics that the other did not share. These characteristics discriminate between the two clinical conditions and are crucial for effective treatment. More in detail, patient SC had the following additional symptoms: A4 "increase or loss of weight", A8 "psychomotor retardation" and A9 "fatigue or energy loss." On the other hand, patient FG, in addition to the shared symptoms, presented the following attributes: A6 "insomnia", A7 "agitation", A15 "suicidal ideation", A19 "irritability", A21 "health concerns" and A22 "somatic disorders".

Furthermore, we considered the two patients' replies to the items of the three sub-dimensions of the QuEDS (cognitive, somatic, and affective). Patient SC responded positively to 12 out of 15 items of the cognitive factor, 11 out of 14 items of the somatic factor, and 7 out of 12 items of the emotional factor; in contrast, FG responded positively to 14 out of 15 items of the cognitive factor, 6 out of 14 items of the somatic factor, and 11 out of 12 items of the affective factor. It has to be stressed, however, that the mere score to the subscales of the QuEDS, even if useful to help clinicians in preliminarily understand the situation of patients, is not sufficient to clearly differentiate the specific kind of depression characterizing the two patients. In fact, neither a high score in the affective factor implies a mixed depression, nor a high score in the somatic factor implies the presence of an inhibited depression. Such characterizations can be, on the contrary, easily deduced by the clinical states provided through the FPA approach (Serra et al., 2017). Indeed, for example, a high score in somatic scale may represent more agitated symptoms, and at the same time more psychomotor retardation. The analysis through FPA allows a better classification of the individual symptoms' case, which therefore allows for planning different pharmacological and psychological treatments for the two patients.

Moreover, the proposed methodology even provides information to clinicians about the possible symptoms of a person who does not have MDEs and who belongs to the non-clinical group. To illustrate, MT obtained a score of six to the questionnaire (the mean of the non-clinical group) after responding affirmatively to items I1, I11, I13, I27, I31, and I33. His clinical state was A3, A9, A13, A20, and A22. Even if he did not show a depressive symptomatology, dysfunctions related to his sexual desire, his energy, and his indecision emerged from his clinical state; also, he has somatic complaints. A usual questionnaire only provides a quantitative score (6), which only means that MT is not suffering from MDEs. This information could underline symptoms in common with some other psychological disorder and could show alarming manifestations, which occur in a "broad spectrum" evaluation, in which the clinician understands some crucial symptoms of the subject and then explores them with more specific and targeted tests (See Chapter 2, CBA 2.0).

Then the main output of the QuEDS is the <u>clinical state</u> of the patient. Indeed, the QuEDS takes into account all of the positive responses of the subject, which are closely linked to the symptoms through the FPA (MDE clinical criteria).

4. *Discussion*

According to the DSM-5, there are different types of MDE, and depressive symptoms may have different features, depending on the individual and his/her particular disorder. The present research was aimed to introducing a new assessment tool capable of account for the differences among the clinical symptomatology of patients that are not evaluable using traditional test scores alone. This task was carried out by using FPA as the theoretical framework in constructing the tool. Concerning the different clinical features, some specific illustrations showed how the tool could be used to more deeply investigate the clinical state of different patients.

Statistical results confirmed the goodness of the proposed tool, in terms of both validity and reliability. The high internal consistency of both the subscales and the whole scale indicates how the items are coherent in exploring the construct. With respect to the test-retest reliability, the correlation shows a good stability of the measure. Concerning the factor structure, the hierarchical model explains best the observed data. As shown in the results MDE can be explained both by the general "depression" factor and by the three sub-factors (cognitive-somatic and affective), and a patient may have more somatic symptoms, or more cognitive/affective symptoms depending on the features of his illness. The results of the divergent/convergent validity on the one hand showed the difference between the correlation of anxiety subscale (A) of the DASS and QuEDS and the correlation between the depression subscale (D) of the DASS and QuEDS, highlighting the convergent validity between the D-scale and QuEDS. On the other hand, the correlation between the subscale stress (S) of the DASS and QuEDS showed the presence of many shared clinical features between the two constructs. Specially, the affective and somatic sub-factors of the QuEDS have a high correlation with the scale S. This result may seem unusual at first, but as the literature suggests (Dumont & Provost, 1999; Hewitt & Flett, 1993; Tafet et al., 2001) stress and depression have many symptoms in common, in particular the people vulnerable to mood disorders are more sensitive to stress (Bidzi, 1984). Furthermore, the Stress scale of DASS is sensitive to levels of chronic non-specific arousal. It assesses difficulty relaxing, nervous arousal, and being easily upset/agitated, irritable/over-reactive and impatient; all these symptoms are highly correlated to the affective dimensions, as our results suggested.

Concerning the content validity, from this new perspective, the relations among the 41 items and the 23 clinical criteria play a crucial role. The matrix shows that all the clinical criteria are investigated by at least one item. This result is very important because it means that the presence or absence of the 23 symptoms selected for

describing the MDE can be detected using the QuEDS. Having information on all of these symptoms makes it possible to compare the observed responses of the subject according to the clinical symptoms he/she demonstrated up to the present. This fact, for instance, allows for determining whether an individual has more inhibited symptoms, more agitated ones, or both. Indeed, the set of items to which the individual responds includes a well-defined series of clinical criteria, which are useful for a first psycho-diagnostic examination. Thus, the output of the QuEDS is no longer crucially related to some sort of cut-off (or score) that shows whether the person could be classified as suffering from the disease or not. On the contrary, it consists of qualitative and quantitative information about the patient's clinical state. Such output provides the potential capability to go beyond the scores and investigate the configuration of the patients' symptoms to differentiate people who received the same score on the test but have different symptoms. Indeed, they responded affirmatively to the same number of items but not to the same items; this allow considering when the whole symptomatology is more serious (Serra et al., 2017).

Moreover, unlike many self-report measures in use to assess MDEs, the QuEDS deeply investigates the symptoms related to agitated depression, including irritability, insomnia, crying spells, somatic disorders, and agitation. The investigation the symptoms related to mixed depression is quite important because people with mixed depression need completely different pharmacological and psychological treatments from those with other types of MDE; for example, antidepressants can increase the psychomotor agitation and the risk of suicide in people with agitated depression (Baldessarini et al., 2006a; Balázs, Benazzi, Rihmer, Rihmer, Akiskal, & Akiskal, 2006).

You can find a similar version of this work in "Serra, Spoto, Ghisi & Vidotto, 2017; Frontiers in psychology." However, in this version, the FPA methodology has not been

repeated since it is already in depth in Chapter 3, and some parts of the introduction and the results in this chapter have been further deepened.

To conclude, another important result obtained by the QuEDS is the possible future application of an adaptive logic based on the matrix in Table 5.6, which will be the core of the next Chapter. Indeed, we divided the matrix of 41 items into three sub-matrices related to the three sub-factors cognitive, somatic, and affective (Table 5.6). This will be useful for the future implementation of the three algorithms of the QuEDS to respond adaptively and individually to the test. The step to achieve this goal will be explain in the next Chapter, the third research.

CHAPTER 6

An adaptive version of the new Major Depressive Episode assessment tool (QuEDS).

1. Introduction and research aim

In clinical practice, self-report questionnaires and interviews are the main methods used for clinical evaluation. As stated in the previous Chapter, the score of a questionnaire is helpful in distinguishing individuals with critical clinical features but is not sufficient to differentiate patients with different symptom configurations who obtained the same score to the test. Moreover, not all the items have the same "weight" from the clinical point of view; in fact, they reflect different symptoms that may be more or less severe (Serra et al., 2017). On the other hand, the main problem of the interviews stands in the possible wrong inferences of the clinician that can lead to misdiagnosis and consequent inappropriate treatment.

Currently, a valid tool that the clinician can use in the assessment is the computerized adaptive testing (CAT; Reise & Waller, 1991; Baek, 1997). In this way, a questionnaire can be administered adaptively, such that an individual responds only to items that are most appropriate for assessing his or her level of impairment mimicking the semi-structured interview (Gibbons et al., 2008). In Chapter 2 the main computerized adaptive assessment methods were described and it was shown how all the mentioned systems, are based on a representation of the domain of interest that is different from that of FPA (as described in the third Chapter). As far as we know, the system we propose is unique as it provides adaptivity, a formal definition of the clinical field, and a probabilistic model to account for both false positive e false negative of each item.

The aim of this work is to develop the adaptive form of QuEDS. We present an adaptive testing system based on a formal representation of items and diagnostic criteria presented in the previous research (Table 5.6), able to provide the adaptive logic of semi-structured interview but in the way of self-report tool. In particular, the three sub-scales of QuEDS (cognitive, affective and somatic) are used in the FPA representation taking advantage of the adaptive assessment algorithm of KST.

1.1. Adaptive assessment through FPA

The three matrices in Table 5.6 (Chapter 5), represent the three clinical contexts through which the three deterministic clinical structures can be obtained. As explained in Chapter 3, the clinical structure may be depicted as lattice that represents the relationships among the items and attributes and among different items. Each node of the lattice associates to each admissible response pattern (K) the set of attributes that an individual endorses if he is in that clinical state. Thus, the clinical structure represents a deterministic "skeleton" that identifies a starting point for adaptive assessment. Given the clinical structure, the deterministic adaptive assessment would proceed as follows:

1. It selects the item that is closest to be present in 50% of the clinical states and proposes it to the patient.

2. It registers the answer (YES-NO) and excludes all the clinical states that do not contain the investigated item if the answer is "yes", or vice versa all the clinical states that contain the state if the answer is "no".

3. It selects the next item by choosing the one closest to be present in 50% of the remaining clinical states;

4. The assessment stops when there is only one clinical state remaining and the output is the clinical state with all the attributes (diagnostic criteria) it investigates.

It is evident how a completely deterministic approach is inadequate for the assessment in clinical practice since it is not representative of reality. The deterministic structure needs to be tested on real data. Indeed, in self-report tools, problems with patient insight or with item wording may produce false positive answers (lucky guess; η) or false negative answers (careless error; β) and not all the clinical states have the same probability to occur. As described in Chapter 3, Section KST, the probabilistic structure (P, K, π) is needed to provide a probabilistic framework to the deterministic structure. The probabilistic model applied is the Basic Local Independent Model (BLIM; Doignon & Falmagne, 1999), which takes into account the probability of false positive (η) and false negative (β) for each item; the probability of occurrence (π) for each clinical state (see pg. 65). Each observed response pattern should derive from the patient's clinical state (latent), the probability of error (η, β) for each item, the probability of the states. Given the clinical structure and the parameters (η, β, π) a probabilistic assessment proceeds as follows:

1. It selects and administers the item that best splits into two equal parts the probability mass of the clinical states (questioning rule).

2. It registers the answer (YES-NO). So, it reduces the probability π of all the states that do not contain the investigated item and increases the probability π of all the other states if the reply is "yes"; vice versa if the answer is "no" (updating rule).

3. It repeats the procedure until one of the states exceeds a predetermined probability value (e.g., .70; stopping rule).

4. The output is the clinical state which represents the most likely symptomatic representation of the patient's situation regarding a specific disorder.

Two studies will be presented in this Chapter. In the first study, the parameters will be estimated on the real data in order to proceed with the adaptive form of the tool. In the

second one, a simulation study is carried out to test the efficacy and accuracy of the implemented adaptive procedure.

2. First study: Parameter estimation

2.1. Method

In order to validate the *three clinical structures* obtained (cognitive, somatic and affective), the parameters of the BLIM have been estimated for each of the three structures based on data from the 383 participants.

Participants: The sample included: a clinical group consisted of 38 subjects with MDE (who were diagnosed with major depressive disorder or bipolar disorder) of the Neurosciences, Mental Health, and Sensory Organ (NESMOS) Department of La Sapienza University, Rome. (This sample is the same of the previous research, chapter 5). A non-clinical sample of 345 individuals recruited in the area of the University of Padova (both students and non-students; 68% were female). The majority of participants had a high school diploma, and their ages ranged between 19 and 58 years. (The criterion of exclusion was the same of the previous research, chapter 5).

Administration: All the research participants completed informed consent and sociodemographic forms before answering the questionnaire items. No time limit was imposed to complete the questionnaires. All 383 subjects completed the written form of QuEDS (41 dichotomous items). This tool has been described in the previous research, presented in Chapter 5. At clinical intake, participants provided written, informed consent for potential research analysis and anonymous reporting of clinical findings in aggregate form, in accord with Italian legal and ethical requirements. The study was conducted in accordance with the Declaration of Helsinki. All participants entered the study of their own free will and they were informed in detail about the aims of the study, the voluntary nature of their participation, and their right to withdraw from the

study at any time. Furthermore, participants could ask for restitution about their own score, providing authors with their own auto generated code, used during the administration phase.

Procedure: The estimate was performed with a specific version of the Expectation-Maximization Algorithm (Dempster, Laird, & Rubin, 1977) for MatLab, i.e., CEMBLIM. For the description of the algorithm, refer to Spoto (2011). The fit of each of the three models has been tested by Pearson's chi-square. It is well known that for large data matrices (as those used in the present study) the asymptotic distribution of χ^2 is not reliable. Therefore, a p-value for χ^2 has been obtained by parametric bootstrap (n. of replications = 5000).

2.2. Results and discussion

In Table 1 are displayed the fit indexes of the three models estimated by CEMBLIM. Results show adequate fit indexes for all three models. In particular, this table shows the global fit indexes obtained for the three models together with the corresponding p-values obtained by parametric bootstrap and the number of clinical states for each sub-scale.

Table 6.1: The global fit indexes of the three models

Sub-factor	Num. of states	*df*	χ^2	Bootstrap P
Cognitive	124	32144	23348	.07
Somatic	163	15972	7237	.16
Affective	142	3928	8696	.06

Table 6.2, displays the false positive (η) and a false negative (β) for each item of the three sub-scales of QuEDS. Note that the sum of η and β do not exceed the value 1 in line with the BLIM model assumptions.

Formally: $\eta + \beta < 1$ for all $q \in Q$ → $\eta < 1- \beta$ and $\beta < 1- \eta$

Indeed, the probability that a false positive (η) occurs must be less than the complement to 1 of the probability of a false negative (β). In other words, the probability of a false positive (η) on an item i, must be less than the answer is "yes" without false positive. Same explanation as regards the false negative.

Table 6.2: Estimated parameters β and η for each item of the three sub-scales

COGNITIVE	β	η	SOMATIC	β	η	AFFECTIVE	β	η
Item 5	0.25	<.001	Item 1	0.20	0.04	Item 7	0.44	0.02
Item 6	0.18	<.001	Item 2	<.001	0.01	Item 8	0.14	<.001
Item 9	0.50	<.001	Item 3	0.13	0.04	Item 12	0.09	<.001
Item 10	0.09	<.001	Item 4	0.01	<.001	Item 15	0.11	0.01
Item 14	0.33	0.01	Item 11	<.001	<.001	Item 17	0.45	0.07
Item 19	<.001	0.03	Item 13	0.15	<.001	Item 18	0.08	0.03
Item 20	0.19	<.001	Item 16	0.12	0.01	Item 29	<.001	<.001
Item 21	<.001	<.001	Item 22	0.10	<.001	Item 34	0.06	0.07
Item 24	<.001	<.001	Item 23	0.05	<.001	Item 36	0.36	0.02
Item 25	0.03	0.01	Item 26	0.02	0.04	Item 37	0.17	0.01
Item 27	<.001	0.18	Item 28	0.06	<.001	Item 38	0.09	<.001
Item 30	0.44	0.01	Item 31	0.12	<.001	Item 40	<.001	<.001
Item 32	0.11	0.06	Item 35	0.16	0.07			
Item 33	0.02	<.001	Item 39	<.001	<.001			
Item 41	0.01	<.001						

As we can see from the table, the single items' η parameters seem quite small for almost all items. On the contrary, there are two items with high probability of false negative (β)

in the cognitive scale of QuEDS: item 9 (β = .50) "I think my life is hell and I only deserve to feel bad", and item 30 (β = .44) "I feel too much on the other people that it would be better if I killed myself". The false negative indicates that some subjects answered "no", although the symptoms investigated by that item were also investigated by other items where individuals answered "yes". This could mean that the interpretation of these specific items could be complicated for some subjects (maybe because they are composed of two sentences) and therefore the symptoms included in those items were more easily understood within other items. To better understand the links among the items and the attributes look at Table 5.5, Chapter 5.

In the affective scale of QuEDS, there are two other items with high false negative rate: item 7 (β = .44) "I keep crying very easily" and item 17 (β = .45) "I often feel like crying, but I cannot do it". Both high values of β are related to "crying"; in this case we could suggest that the feel like crying as well as the "crying" were underestimated in the non-clinical sample. We can suppose that the subject could either intentionally fake the specific answer, or the subject's answer could be affected by the poor introspection about "crying". However, all the other items show reasonable error parameters.

Parameter estimates will be essential for the implementation of the algorithm to obtain the adaptive form of the QuEDS.

3. *Second study: Experimental simulation*

3.1. Introduction

The aim of this simulation study was reproducing the clinical testing in an adaptive way using real data. This study presents the adaptive assessment algorithm for the three QuEDS sub-scales (cognitive, somatic, affective), which is an extension and adaptation to the clinical domain of an algorithm designed for the Adaptive Knowledge Assessment of an individual (AKA algorithm; Falmagne & Doignon, 2011) namely the

ATS-PD (Donadello et al., 2016). In the present section, we used all the concepts described in Chapter 3, and explain the operating principles of the ATS-PD algorithm.

3.2. ATS-PD Algorithm

The algorithm is based on three main steps: the questioning rule, the updating rule, and the stopping rule. The *questioning rule* selects the item to ask, i.e. the item q ''maximally informative'', such that the sum of the likelihoods of all the states containing q has to be as close as possible to .50. If several items are equally informative, one of them is chosen at random. At each step of the procedure, we call $L_n(K)$ the probability of the state K at the step n. The subject's response is collected by the system, then the *updating rule* is applied to obtain the likelihood $L_{n+1}(K)$ of every state K. If a subject answers "yes" to q, the likelihoods of the states containing q are increased and, correspondingly, the likelihoods of the states not containing q are decreased. On the contrary, if a subject answers "no" the likelihoods of the states containing q are decreased and, correspondingly, the likelihoods of the states not containing q are increased. If we indicate an affirmative response with r = 1 and a negative one with r = 0, we can formalize the updating rule as:

$$L_{n+1}(K) = \frac{\zeta^K L_n(K)}{\sum_{K' \in \mathbf{K}} \zeta^K L_n(K')}$$

Where:

$$\zeta_{q,r}^K \begin{cases} \zeta_{q,1} & if\ q \in K, r = 1 \\ 1 & if\ q \notin K, r = 1 \\ 1 & if\ q \in K, r = 0 \\ \zeta_{q,0} & if\ q \notin K, r = 0 \end{cases}$$

In which $\boldsymbol{\zeta}$ is a parameter that increases the likelihood and influences the efficiency of the adaptive assessment process, and $\boldsymbol{\zeta}$ > 1. The higher is the value of $\boldsymbol{\zeta}$, the more

reliable are considered the answers provided by the subject. It has been observed (Falmagne & Doignon, 2011) that ζ values less than 2 make the assessment redundant since a higher number of answers are needed to achieve a reliable conclusion about the patient' clinical state K. However, assign an high fixed value to ζ could affect algorithm efficacy because it can generate errors due to the overestimation of the truth of the answer, so that the output of the assessment can be returned in a very fast but not exact way. It has been proven that the algorithm, with a value of 21 of the parameter, should tend to converge to the correct clinical state in different applications, e.g. ALEKS (Falmagne & Doignon, 2011). An alternative way to estimate ζ is to assign the ζ value based on the η, β parameters according to Koppen (see Doignon & Falmagne, 2011, pg. 265). Formally:

$$\zeta_{q,1} = \frac{1-\beta_q}{\eta_q} \quad \text{and} \quad \zeta_{q,0} = \frac{1-\eta_q}{\beta_q}$$

The rule based on η, β is local, meaning it takes into account the false positive and false negative of the last item asked to the individual. If the answer is "yes" then the value of ζ is calculated according to the first formula, otherwise if the answer is "no", the value is calculated according to the second formula. In this way, the parameter ζ is conditioned by the probabilities of false positive and false negative, so that if the answer is "yes" and if η is a high value, the ζ value will be always lower, and the answer will have a lower weight in the assessment. If the answer is "no", and if the value of β is high, at equal measure, the ζ value will be lower and the answer will have a lower weight in the assessment (see Koppen, Doignon at Falmagne, 2011, pg. 265).

At this point of the procedure, the algorithm performs the *Bayesian updating rule* (Donadello et al., 2016). Formally:

$$P(K_i|R) = \frac{P(R|K_i)\boldsymbol{L}_n(K_i)}{\sum_{j=1}^{|K|} P(R|K_j)\boldsymbol{L}_n(K_j)}$$

where $P(R|K_i) = p(\text{R}, \text{K}_i)$ is given by Equation:

$$p(R) = \sum_{K \in \mathcal{K}} \rho(R, K)\pi(K)$$

The algorithm stops when a *stopping condition* is satisfied. It means that the procedure is repeated with the questioning rule and the updating rule until one of the states exceeds a predetermined probability value fixed to .70. The outcome of the assessment is the *clinical state* that reaches this predetermined high probability. At this point, the attributes that are implicated in this state with their relative probabilities are displayed.

3.3. Simulation Design

Six different conditions were considered in which the following two variables were manipulated:

- ✓ The value of the parameter **ζ** used for the updating rule: Two levels were defined for its value; in the first situation, it was set to 21; in the second one **ζ** value was calculated on the basis of η, β as shown above.

- ✓ The Bayesian updating rule (see Donadello et al., 2016): Three levels were defined for this variable: online, offline, absent. When the Bayesian updating is "online", it means that the subject's answers for each item asked, step by step, were taken into account to update the likelihood of each state in each step. If we use the Bayesian updating rule "offline" the update is implemented only at the end of the simulation. Finally, when the Bayesian updating is "absent" means that this rule is not used at no stage of the simulation.

Table 6.3 contains the six possible conditions, obtained by manipulating the two variables as descripted above.

	Parameter ζ	Bayesian update
1	21	offline
2	21	online
3	21	absent
4	μ, β function	offline
5	μ, β function	online
6	μ, β function	absent

3.4. Method

Participants: The sample is the same of the previous study: 383 subjects (38 clinical-sample, 345 non-clinical sample).

Procedure and administration: The ATS-PD algorithm takes as input the clinical structure of each sub-scale of QuEDS (cognitive, somatic, affective) and the parameters estimates (the probability (η, β) of false positive and false negative for each item and the probability of the states K). Thus starting from a probabilistic structure (P, K, π), it is possible to assign to each response pattern (R) its probability p(R, K) given a knowledge state K. Formally:

$$p(R) = \sum_{K \in \mathcal{K}} \rho(R, K)\pi(K)$$

Because the response function satisfies local independence for each item $q \in Q$, the conditional probability p(R, K) is determined by the probabilities β, η related to each item q. Formally:

$$\rho(R, K) = \left[\prod_{q \in K \setminus R} \beta_q\right]\left[\prod_{q \in K \cap R} (1 - \beta_q)\right]\left[\prod_{q \in R \setminus K} \eta_q\right]\left[\prod_{q \in \overline{K \cup R}} (1 - \eta_q)\right]$$

These last equations reported represent the basic local independence model (BLIM; Doignon & Famagne, 1999). Once the error parameters were estimated (previous study), the algorithm (ATS-PD) was applied for the adaptive simulation of the written form of QuEDS sub-scales.

The clinical assessment phase was reproduced, evaluating whether the adaptive form of QuEDS could generate the same response pattern of a subject who answered the written version (system efficacy), but with a reduced number of questions (system efficiency). The reproduction of the testing started when the system imports the clinical structures of the subscales and the response patterns of 383 subjects. For every response pattern R, the system performs the testing by asking a question, answering automatically, and updating the likelihood until it uncovers the latent state K. Since the number of the states in the structures was relatively high compared to the sample size, the estimates of π_k were not reliable. Therefore, we choose to use a *uniform probability distribution* on all the states $L_n(K)$ at time 0.

To better understand the results of the simulation we introduce some concepts. A response pattern R is a list of the observed answers provided by a subject in the written version of QuEDS. We have the *response patterns (R_1, R_2, R_3 ….R_n)* of the 383 participants. We call K* the state in the structure $\mathcal{K}$ that is the output of the adaptive assessment given a response pattern R_n. It is important to emphasize that the output of the assessment for any response pattern R_n through the adaptive simulation is always a state of the clinical structure. We call d(K*, R) the cardinality of the set difference between the output K* and the pattern R.

The results of the simulation for each subject can fall into one of the following categories:

1. K*∈ $\mathcal{K}$ = R. This happens when the response pattern R is a state K of the clinical structure $\mathcal{K}$. In this case the distance between the state and the response

pattern R is: d(K*, R) = 0. In the specific condition, the output of the adaptive assessment is exactly the same of the output obtained with the written version of the sub-scale.

2. K* ∈ 𝒦 ≠ R In this case of course, d(K*, R) > 0. It means that the response pattern R is not a clinical state K of the structure 𝒦. In this situation two alternatives may occur: in the first case the distance d(K*, R) is minimum. It means that there is no K ∈ 𝒦 such that d(K, R) < d(K*,R). In the second case the distance d(K*, R) > minimum. It means that exist K ∈ 𝒦 such that d(K, R) < d(K*, R). indeed, there is another clinical state K in the structure that is closer to the response pattern R even if it is not the output of the assessment. This latter situation can occur for several reasons; it may depend on the error parameters (η, β), on the sequence of questions posed by the system, and on the Bayesian update and parameters used by the algorithm.

3.5.Results and Discussion

The aim of this study was reproducing the clinical testing in an adaptive way using real data (N =383) to verify if:

✓ The new system of QuEDS is able to identify as output of the assessment the clinical states corresponding to the response patterns (*accuracy*).

✓ The adaptive form of QuEDS asks a smaller number of questions with respect to the standard written version questionnaires (*efficiency*).

As we know, the adaptive form of QuEDS sub-scales return only clinical states in the clinical structure. We verified that the three clinical structures are good model of the reality, after fit indexes estimates (Spoto et al., 2010). Thus, a response pattern R, with an assigned clinical state K in which d(K*,R) ≠ 0 , indicates that R is affected by false-positive or false-negative errors.

In our clinical structure, the simulation works better with the Bayesian updating rule online, and with the parameter **ζ** calculated on the basis of η, β.

Tables 6.4, 6.5, 6.6 display the comparison of the main indexes of the different tested adaptive procedures. Each table relates to one of the three QuEDS subscales (cognitive, somatic, affective). In particular, the first column contains the two solutions (explained in the simulation design section) in which was assigned the value of parameter **ζ.** The second column refers to the three ways in which we manipulated the Bayesian update (online, offline, and absent). The third column contains the maximum number of questions asked to finish adaptive assessment and reach the output. The fourth column contains the number of the R (response pattern) in which d(K*, R) > minimum. Thus, in these cases the response pattern R was not a state K of the structure, and the output of the assessment is not the state of the clinical structure closest to R.

Table 6.4: Cognitive factor

parameter ζ	Bayesian update	n. max of questions	n. states d (K,R) ≠ 0 d(K,R) ≠ minimum
21	offline	17	12
21	online	11	15
21	absent	17	11
μ, β function	offline	20	7
μ, β function	**online**	**11**	**10**
μ, β function	absent	20	7

Table 6.5: Somatic Factor

parameter ζ	Bayesian update	n. max of questions	n. states d (K,R) ≠ 0 d(K,R) ≠ minimum
21	offline	9	8
21	online	9	6
21	absent	9	8
μ, β function	offline	11	8
μ, β function	**online**	**9**	**5**
μ, β function	absent	11	8

Table 6.6: Affective factor

parameter ζ	Bayesian update	n. max of questions	n. states d (K,R) ≠ 0 d(K,R) ≠ minimum
21	offline	8	40
21	online	8	35
21	absent	8	39
μ, β function	offline	23	39
μ, β function	**online**	**8**	**29**
μ, β function	absent	29	34

As we can see from the tables in all three clinical structures, the system works better with *online Bayesian update* and, *with the parameter ζ calculated on the basis of η, β*.

Starting from the cognitive scale, which has 15 items in total, with this condition we have a maximum of 11 questions asked and a minimum of 7 question (items) asked in the adaptive form to achieve the output of assessment; the average is 8,83 items asked (s.d= .47). It means that the saving in terms of question posed is between 31% and 53% for the cognitive factor of QuEDS. We found 10 response patterns R in which the condition d(K*, R) = minimum was not satisfied, so K* is not a minimum distance to a response pattern R. We have in total 383 subjects and 129 different response patterns; 10 of the 129 response patterns R in writing form were not a minimum distance. It means that the states K* generated as outputs by the algorithm are states of the clinical structure that are not as close as possible to the response patterns R; indeed, there are some clinical states in the structure that are closer than K* to each of the 10 response patterns R even if they are not the output of the assessment. Formally we say that:

$$d(K, R) < d(K^*, R)$$

In the specific case, d(K, R) - d(K*, R) ≤ 2. It means that the distance between the state K* output of the assessment and the state closer to the response pattern is never greater than 2.

The somatic scale has 14 items in total. Through the Bayesian update online simulation and the parameter ζ calculated on the basis of η, β, we have a maximum of 9 items asked and a minimum of 8 item asked to achieve the output of the assessment; the average is 8,42 items asked (s.d= .82). It means that the saving in terms of questions posed is between 36% and 50% for the somatic factor of QuEDS. We have in total 173 different response patterns, and we found 5 response patterns R in which the condition d(K*, R) = minimum, was not satisfied. There are some clinical states closer than the clinical state K*, output of the assessment, for those 5 response patterns.

In the specific case, d(K, R) - d(K*, R) ≤ 2. It means that the distance between the state K* output of the assessment and the state closer to the response pattern is never greater than 2.

Finally, on the affective scale, some difficulties emerged. Indeed, we have a total 12 items in the written version and using the same condition, we have a maximum of 8 item asked and a minimum of 7 items asked in the adaptive version (experimental simulation); the average is 7,66 items asked (sd= .47). The saving in terms of questions posed is between 33% and 42% that is on average lesser than the saving in the other two scales. Moreover we found 29 response pattern in which the condition d(K*, R) = minimum was not satisfied (as defined above). In the specific case d(K, R) - d(K*, R) ≤ 3. It means that the distance between the state K* output of the assessment and the state closer to the response pattern is never greater than 3.

4. *General Discussion*

The adaptive form of QuEDS is based on an extension of an algorithm for the assessment of knowledge (Falmagne & Doignon, 2011). Several new features are added to the procedure according to Donadello et al. (2016): first, the definition of the reference structure for the algorithm is performed through the application of FPA.

Second, the parameters estimates are carried out by referring to CEMBLIM (Dempster, Laird, & Rubin, 1977). Third, the parameter **ζ** was calculated on the basis of η, β; finally the algorithm updates the states' probabilities through a Bayesian rule step by step online in the testing.

The aim of this study was to build an adaptive version of QuEDS using its sub-factors (cognitive, somatic, affective). This tool should support the clinician in the diagnosis of different types of major depressive episode. Indeed, as explained in the previous Chapter, the numerical score provides information on the level of depression but is not able to differentiate individuals with the same score but different symptoms. The FPA through the relationship between items and symptoms allows to obtain qualitative information. However, with the written version, the clinician's work to obtain the information is time-consuming, and errors in interpreting the answers may also occur as the error parameter estimates are not considered.

In the first study, in order to validate the obtained three clinical structures (cognitive, somatic and affective), the parameters of the BLIM have been estimated for each of the three structures based on data from the 383 participants. The models showed adequate fit indexes; the estimated μ and β were essential for the implementation of the algorithm in order to obtain the adaptive form of the QuEDS.

The ATS-PD algorithm taking into account the false positive and false negative answers, allows to individualize a patient's precise critical areas in efficient way, that is, a system poses fewer questions than the standard written questionnaire does. Moreover, it is able to achieve a clinical evaluation that goes beyond the score, indicating, adaptively, all the problems presented by an individual in a way that mimic the semi-structured interview.

The performance results reported in the previous section showed that all of the response patterns that are states are assigned to that state, so the system is able to correctly

reproduce the patient's admissible response pattern of a questionnaire. It is worth noting some crucial differences between a classical questionnaire administration and the assessment through ATS-PD algorithm.

The adaptive form of QuEDS returns the individual's clinical state for each sub-scale, with the subset of symptoms included. Indeed, the clinical state is the subset of items in which the individual answers affirmatively with a set of attributes (in terms of symptoms) endorsed by those items. This information permits to distinguish the individuals showing different critical areas and, thus, leading to different diagnoses. In Chapter 5 we showed the importance of obtaining the clinical state for differential diagnosis. The adaptive model allows to reach the clinical state as a test output directly without the clinician's hard work.

According to the model (supported by the fit indexes) the admissible response patterns should be the states of the structure, so the system will always complete its evaluation in one of these states. However, the response patterns observed could be affected by errors (μ, β) and could differ from the assigned state. Indeed, the response patterns that are not states are assigned to states with $d(K^*, R) = $ minimum.

In some cases, we have seen that $d(K^*, R) > $ minimum. This means that some response patterns R (as input) had a clinical state K^* as output that was not as close as possible to R. Indeed, there was another K state of the structure closer to the pattern. This could depend on the sequence of questions asked by the system, either by the error parameters (μ, β), or by the type of update used by the system. However, this situation has rarely occurred in the cognitive and affective sub-scales, while in the case of the affective sub-scale it has occurred more often.

Another important result to be stressed is the reduction of the number of questions asked together with the improvement of the quality and quantity of information collected. In the classical written form of the QuEDS, each participant has to answer all

41 items, 15 of the cognitive sub-scale, 14 of the somatic subscale, 12 of the affective subscale. In the adaptive form of QuEDS only a percentage ranging between 50% and 70% of the items are asked and the clinician is provided with the clinical state of the individual, including the diagnostic symptoms presented by the patient.

Thus, the adaptive version of QuEDS differentiates the individual's depressive symptoms beyond the score and allows to administer only the items related to its symptomatology following the logical flow of question-answer.

As future work, we intend to extend the number of participants in order to obtain more precise fit indexes, and more reliable error parameters. Another objective is to achieve a simple graphical user interface providing the clinician with a helpful way to interact with the system. Finally, there can be several improvements of ATS-PD system, for example the possibility of simplifying the updating rule for real-time application of QuEDS (i.e. Augustin, Hockemeyer, Kickmeier-Rust, Podbregar, Suck, & Albert, 2013). Another important future direction will be to extend the adaptive version of the tool also with politomicous items with Likert scale. There are two main solutions under evaluation to solve this issue: the fuzzy logic approach and an IRT oriented solution. In both cases, there is the possibility to take into account the case in which the answering format is not dichotomous. The implementation of either of these proposals would allow FPA to construct tools with not only dichotomous replies works becoming much more fruitful in clinical practice.

CHAPTER 7

Clinical evidence for *differential diagnosis* of Agitated Depression

1. Introduction and research aim

In the previous Chapters, we demonstrated the possibility of achieving a differential diagnosis of depressive symptoms starting from a new evaluation approach. Through the FPA, we built a tool that is able to differentiate patients with the same score on the questionnaire but with different symptom configurations. We have also shown how a computerized adaptive system can mimic the semi-structured interview process in which only the patient's symptomatic areas are investigated.

In evaluating depression, an adaptive tool able to going beyond the numerical score is essential. In fact, clinicians have to face different types of major depressive episodes that often requiring a different diagnosis and a specific treatment. The case of agitated depression, or mixed, represents perhaps the most important. In fact, agitated depression is classified as a mixed episode, with both depressive and manic symptoms, and for this reason it can not be treated in the same way as the typical depressive episode.

The research that will be presented below has been carried out in collaboration with the University of Cardiff and Worcester, in particular with the Bipolar Disorder Research Network (BDRN) during the research period abroad.

In line with the topics discussed in the previous Chapters regarding the importance of differential diagnosis, a study of 3750 patients with bipolar disorder or major depression has been conducted with the aim of deepening all issues related to agitated depression.

The study was part of the ongoing programme of research into the genetic and non-genetic determinants of BD and related mood disorders (Bipolar Disorder Research Network, BDRN; bdrn.org) which has UK National Health Service (NHS) Research

Ethics Committee approval and local Research and Development approval in all participating NHS Trusts/Health Boards. Data come from BDRN.

Agitated Depression (AD) still have an unclear place in mood disorders and different definitions (Akiskal et al., 2005; Benazzi et al., 2002; Koukopoulos, & Sani, 2014) despite it is not at all a rare observation and many authors underscored the necessity of a corrected diagnosis in order to avoid erroneous treatment (Baldessarini et al., 2006a, 2006b; Olin et al., 2012; Vázquez et al., 2013). Kraepelin (1913; 1921) classified agitated depression as a result from the combination of opposite polarity of symptoms: mood and thought in depressive polarity and activity in manic polarity. In his view, it was enough to have one of the three components (psychomotor activity, mood and thinking) in manic polarity to have mixed state (Akiskal & Benazzi, 2004). According to DSM-5 criteria (American Psychiatric Association, 2013), Major Depressive Episode (MDE) with mixed features replaced the agitated depression of Research Diagnostic Criteria (RDC, Spitzer, Endicott, & Robins, 1978) and it is defined as a depression with at least three manic/hypomanic symptoms. However different authors consider the DSM-5 picture of AD as a very rare feature with the lack of satisfaction of these criteria (Faedda et al., 2015; Koukopoulos, & Sani, 2014; Maj et al., 2003). Koukopoulos and colleagues (1999; 2014) defined AD as an MDE with at least one of the following criteria: inner psychic tension, psychomotor agitation, irritability, and racing/crowded thoughts. Olgiati, Serretti, & Colombo (2006) defined AD as MDE with OPCRIT item of agitated activity (excessive repetitive activity, such as fidgety restlessness, wringing of hands, pacing up and down) all usually accompanied by expression of mental anguish (McGuffin et al., 1991). There is therefore no univocal definition of agitated depression.

Regarding the ratio of AD in Major Depressive Disorder (MDD), Bipolar Disorder I (BD-I) and Bipolar Disorder II (BD-II), different findings have been reported although

currently the classification of AD in mood disorder's population is still unclear. In Bipolar I reports of AD vary from 20% (Maj et al., 2003) to 50% (Koukopulus et al., 2007). In Bipolar II reports of AD vary from 30% (Benazzi et al., 2002) to 52% (Takeshima, & Oka, 2013). In MDD reports of AD vary from 11% (Benazzi et al., 2002) to 30% (Koukopulus et al., 2007). The various ways of defining and assess agitated depression in many different studies (Akiskal et al., 2005; Benazzi 2004a; Koukopulus et al., 2007; Maj et al., 2003; McIntyre et al., 2015; Takeshima & Oka 2013; Olgiati et al., 2006) could mirror the different findings found on prevalence in MDD, BD-I, and BD-II.

The previous studies on AD have been performed in a relatively small sample (Benazzi et al., 2002; Koukopoulus et al., 2007; McIntyre et al., 2015 etc.); moreover only few studies were focused on the prevalence of AD in Bipolar Disorder I and II and in the differences of the lifetime and episode features linked with AD in the two disorders. Indeed the studies focused more on exploring illness course and hypomanic features related to AD in MDD to investigate possible indicators of bipolarity as well as mood switching in the polarity of episode induced by antidepressant drugs (Akiskal et al., 2005; Angst, Gamma, Benazzi, Ajdacic, & Rössler 2009; Biondi, Picardi, Pasquini, Gaetano, & Pancheri 2005; Olgiati et al., 2006).

Concerning the possible lifetime features associated with AD, some studies correlated the AD with lower age at onset (Benazzi et al., 2002; Koukopoulus et al. 2007) of the mood disorder, bipolar spectrum (Akiskal et al., 2005; Sato et al., 2003), longer duration of the illness period (Maj et al., 2003; Benazzi et al., 2002). Agitated episode was very often associated with intra-depressive hypomanic symptoms (Maj et al., 2003; Oligiati et al., 2006; Perugi et al., 2001), more severity of depressive symptoms (Benazzi et al., 2002; Oligiati et al., 2006), more atypical features specifiers such as mood lability (Benazzi et al., 2002). In Bipolar disorder I patients with AD compared to patients with

non-AD were more likely to receive standard antipsychotic drugs during that episode (Goodwin & Jamison, 2007; Maj et al., 2003). The demographic features correlates with AD according to the previous studies are female gender (Benazzi et al., 2002; Koukopulus et al., 2007; Maj et al., 2003), lower educational levels (Oligiati et al., 2006), and more family history of bipolar disorders (Akiskal et al., 2005; Koukopulus et al 2007; Maj et al., 2003). Moreover, psychomotor agitation and suicidal ideation were found to be correlated in many studies (Andreasen and Grove, 1982; Kendler et al., 1996; Korszun et al., 2004; Maj et al., 2003; Raskin et al., 1969; Sullivan et al., 2002).

Agitated depression have a significant clinical relevance, nevertheless, there are limits of information, underestimation of the consequences, which could result in misdiagnosis of AD and inappropriate/wrong treatment, often with very dangerous outcomes (Akiskal et al., 2005; Bocquier et al., 2013). In particular, treatment with antidepressant drugs of AD could worsen the excitatory symptoms with the failure to relieve the patient's pain (Baldessarini et al., 2006; Koukopoulos, & Koukopoulos, 1999; Vázquez et al., 2013). Moreover, antidepressants monotherapy in AD could increase the risk of suicide (Akiskal et al., 2005; Baldessarini et al., 2006a; Koukopoulos, & Koukopoulos, 1999; Vázquez et al., 2013).

In keeping with this background, the present study have two main aims:

The first one concerns the re-evaluation of the prevalence of AD in the worst episode in a large sample (N= 3750) of patients with MDD, BD-I and BD-II exploring the differences in the rate of AD among the three disorders.

The second aim is to investigate the possible correlations of AD with socio-demographic variables, lifetime features, and the episode features in line with the other researches. Unlike the previous studies, we focus on the differences in Bipolar I and Bipolar II (BD-I, BD-II) agitated depression because many researches have already focused on AD in unipolar disorder (MDD) and treatment-related issues (Biondi et al,

2005; Akiskal et al., 2005; Oligiati et al., 2006; Swann, 2013). Moreover, previous studies have highlighted the association between psychomotor agitation and risk for mood-switching in MDD, suggesting that agitated depression was almost always associated with indicators of bipolarity (Akiskal et al., 2005; Angst et al., 2009; Iwanami et al., 2015) and, MDD patients are often reclassified in the bipolar spectrum (Benazzi, 2006; Cassano et al., 2012; Fiedorowicz, Endicott, Leon, Solomon, Keller, & Coryell, 2011). Thus, to avoid confusion, we preferred to focus on patients with Bipolar Diagnosis.

2. *Materials and methods*

2.1.Participants

Participants are recruited systematically through NHS psychiatric services and non-systematically using advertisements for volunteers via the BDRN website, leaflets, posters, media coverage about the research and also through the UK-based user-led charity, Bipolar UK. All patients in the UK who have a diagnosis of bipolar disorder I, II or major depressive disorder, and are aged 18 years or over are eligible to take part in the BDRN study and enrol after giving written informed consent. The exclusion criteria are patients who: a) have only experienced affective illness as a consequence of alcohol or substance abuse or dependence; b) have only experienced affective illness as a consequence of medical illness or medication; c) have an organic brain disorder or other cognitive problem that impedes their ability to complete the assessments; d) are biologically related to another study participant.

Participants in the current study (N= 3750) were those from whom we had information on the presence or absence of *agitated features* during a depressive episode who met DSM-IV diagnostic criteria for BD-I (N= 2123), BD-II (N= 915), and MDD (N= 712).

2.2. Psychiatric Assessment

Clinical data for each individual in the BDRN study is collected by a trained BDRN interviewer (research psychologist or psychiatrist) using a semi-structured psychiatric interview, the Schedules for Clinical Assessment in Neuropsychiatry (SCAN, Wing et al., 1990). Further clinical data are gathered from participants' psychiatric case notes. Diagnoses and clinical ratings are made using all available clinical data according to pre-specified guidelines. The OPCRIT criteria are used to evaluate the worst depressive episode of participants (McGuffin et al., 1991). Diagnostic and clinical ratings are made by at least two members of the research team blind to each other's ratings. Inter-rater reliability was formally assessed using 20 random cases. Mean kappa statistics were 0.85 for DSM–IV diagnoses and ranged between 0.81 and 0.99 for other key clinical categorical variables. Mean intra-class correlation coefficients were between 0.91 and 0.97 for key clinical continuous variables. Team members involved in the interview, rating and diagnostic procedures were all research psychologists or psychiatrists.

2.3. Agitated Depression (AD) definition

We defined AD according to the OPCRIT (Operational Criteria for Psychotic Illness checklist) definition, which requires the presence of excessive repetitive activity (such as restlessness, wringing of hands, pacing up and down) all usually accompanied by expression of mental anguish. In this study, agitated activity was rated as present/absent during the worst ever episode of depression and mean kappa is 0.85.

2.4. Data analysis

The data were analysed using SPSS (2012). The majority of data were normally distributed so parametric statistical tests were used. The ratio of Agitated Depression present/absent in the three different disorders (BD-I, BD-II, and MDD) was compared using chi-square test. Demographic and clinical characteristics of the two groups (AD

and non-AD) in BD-I and BD-II were compared using chi-square tests for categorical data and t student test for independent samples for continuous data. All OPCRIT episodic features significant at P ≤ 0.05 in the univariate analyses were included as explanatory variables in logistic regression models using the enter likelihood ratio method with absence or presence of AD as the outcome/dependent variable. All the demographic and lifetime characteristics significant at P ≤ 0.05 in the univariate analyses were included as predictors in a logistic regression models using the enter likelihood ratio method with absence or presence of AD as the outcome/ dependent variable.

3. Results

3.1. The prevalence of Agitated Depression in MDD, BD-I and BD-II

There are no significant differences in the demographic features among MDD, BD-I and BD-II. The people with agitated features are 1098 (29,3%) out of 3750.

There is a significant greater proportion of AD in BD-II (see Table 1): 37% of the sample; 29% in BD-I, and 21% in MDD. The difference in the proportion of AD in the different Diagnosis is significant (χ^2= 46.407, p<.001). In particular, all the different ratios are significantly different. AD in BD-I and BD-II (29% vs. 37%, p<.001), AD in MDD and BD-I (21% vs. 29%, p<.001), AD in MDD and BD-II (21% vs 37% p<.001).

TABLE 7.1: Agitated depression ratios in BD-I, BD-II and MDD

DSM DIAGNOSIS	NO Agitated features	YES Agitated features	tot
BD-I	1513 (71%)	610 **(29%)**	2123
BD-II	579 (63%)	336 **(37%)**	915
MDD	560 (79%)	152 **(21%)**	712
Tot	2652 (71%)	1098 (29%)	3750

3.2. Demographic and lifetime features according to agitated depression.

There wasn't significant difference in the proportion of male and female, both in BD-I that in BD-II. There was a significantly greater proportion of lowest educational attainment in AD compared with no AD in patients with BD-I and II.

Concerning the lifetime characteristics related to agitated features in BD-I (see Table 7.2), the individuals in the agitated group had significantly:

- Higher rate of history of alcohol misuse, defined as > 14 units in women and >21 units in men at any point of life, with related impairments (50% vs. 42%).

- Higher rate of history of panic attacks (74% vs. 54%).

- Higher rate of dysphoric mania episodes, defined as manic episode in which the predominant state was dysphoria- i.e. an unpleasant state characterized by unease or mental discomfort including low mood (52% vs. 36%).

- Higher rate of history of suicide attempts (59% vs. 44%).

- Higher rate of rapid cycling, defined as 4 or more affective episodes in one year (37% vs. 28%).

- Higher rate of depressive polarity as the first episode (81% vs. 66%).

- Younger age at illness onset, defined as the age in which symptoms of affective disorder first caused significant impairment (22 vs. 24 years).

The lifetime characteristics related to agitated features in BD-II showed that agitated group had significantly:

- Higher rate of history of Alcohol misuse, defined as > 14 units in women and >21 units in men at any point of life, with related impairments (58% vs. 48%).

- Higher rate of history of panic attacks (85% vs. 64%).

- Higher rate of dysphoric (Hypo)mania episodes, defined as hypomanic episode in which the predominant state was dysphoria- i.e. an unpleasant state characterized by unease or mental discomfort including low mood (33% vs. 25%).

- Higher rate of history of suicide attempts (56% vs. 47%).

- Higher rate of rapid cycling, defined as 4 or more affective episodes in one year (54% vs. 44%).

There were no other significant differences in BD-II lifetime characteristics between the agitated and no agitated group.

TABLE 7.2. Demographic characteristics, and lifetime features according with agitated features in the worst episode of depression.

DEMOGRAPHIC characteristics	BIPOLAR DISORDER I AGITATED FEATURES			BIPOLAR DISORDER II AGITETED FEATURES		
	NO (1513)	YES (610)	χ²(df=1) p	NO (579)	YES (336)	X² (df=1) p
Sex, n (%) **MALE** **FAMALE**	496 (33%) 1017 (67%)	179 (29%) 431 (71%)	χ²=2,370 p=.124	188 (33%) 391 (67%)	100 (30%) 236 (70%)	χ²=.723 p=.395
Education, n (%) **Non graduate** **Degree/Postgrad Degree**	804 (57%) 606 (43%)	365 (65%) 198 (35%)	χ²=10,164 **p=.001**	320 (57%) 238 (43%)	219 (69%) 77 (31%)	χ²=10,950 **p=.001**
LIFETIME FEATURES						
Alcohol abuse	586 (42%)	288 (51%)	χ²=11,96 **p=.001**	269 (48%)	186 (58%)	χ²=8,01 **p=.005**
Cannabinoids abuse	296 (20%)	133 (23%)	χ²=1,52 p=.217	149 (26%)	101 (31%)	χ²=2,60 p=.107
Panic comorbidity	419 (54%)	278 (74%)	χ²=42,25 **p<.001**	244 (64%)	204 (85%)	χ²=31,57 **p<.001**
Dysphoric Mania	461 (36%)	263 (52%)	χ²=37,98 **p<.001**	121 (25%)	86 (33%)	χ²=4,09 **p=.042**
Suicide attempts	655 (44%)	349 (59%)	χ²=37,09 **p<.001**	284 (47%)	185 (56%)	χ²=3,69 **P=.040**
Rapid Cycling	309 (28%)	148 (37%)	χ²=11,23 **p=.001**	161 (44%)	99 (54%)	χ²=7,52 **P=.006**
Polarity 1st episode **DEPRESSION** **MANIA**	853 (66%) 439 (34%)	404 (81%) 98 (20%)	χ²=36,54 **p<.001**	465 (92%) 41 (8%)	268 (92%) 23 (8%)	χ²=.59 p=.75
Impairment age at onset **MEAN** **S.D** **RANGE**	24 9,88 4-73	22 8,70 4-55	T=-3.674 **p<.001**	21 9,15 5-58	20 8,80 5-63	T=-1,89 p=.059

Total vary due to unknown/missing data

The demographic and lifetime characteristics that were significantly higher in the agitated group (p≤ .05) of BD-I were included in the logistic regression analysis. The characteristics that best predicted the presence of agitation in BD-I group were history of panic attacks (OR 1.932, 95% C.I: 1.271 – 2.936, p=.002); lifetime episodes of dysphoric mania (OR 1.689, 95% C.I: 1.140 – 2501, p=.009); lower educational attainment (OR .661, 95% C.I: .448-.976, p=.037) and depression as the polarity of the first episode (OR .584, 95% C.I: .360 – .948, p=.029). This model considers 15,4% of the variance and correctly classified 72% of participants who have agitated features or not.

At the same way of BD-I, the demographic and lifetime characteristics that were significantly higher in the agitated group of BD-II (p ≤.05) were included in the logistic regression analysis. The characteristics that best predicted the presence of agitation in BD-II group were history of panic attacks (OR 2.655, 1.416 – 4.978, p=.002) and suicide attempts (OR 1.919, 95% C.I: 1.129 – 3.264, p=.016). This model considers 10% of the variance and correctly classified 66,3% of participants who have agitated features or not.

Furthermore, in the Agitated group there was a greater lifetime use of psychiatric drugs. In particular BD-I individuals had a significant higher use of antidepressant (96% vs. 88%, p<.001), anxiolytics (76% vs. 64% p<.001), and hypnotics (78% vs. 71%, p=.004); BD-II individuals had a significant higher use of anxiolytics (70% vs 55%, p<.001) and antipsychotics (66% vs 55%, p=.004).

3.3.Worst episode symptoms according to agitated features.

In BD-I almost all the symptoms of the worst depressive episode are more connected with agitated features (see Table 7.3). Indeed, there were significantly greater

proportions of depressive symptoms in the agitated group compared to no agitated group (Loss of pleasure, diurnal variation, suicidal ideation, etc.).

Even in BD-II disorder, many severe symptoms in the worst depressive episode are associated with agitated features (see Table 7.3).

TABLE 7.3: Worst episode symptoms associated with agitated features

OPCRIT VARIABLES	BIPOLAR DISORDER I AGITATED FEATURES			BIPOLAR DISORDER II AGITETED FEATURES		
N (%)	NO (1513)	YES (610)	χ^2 (df=1)p	NO (579)	YES (336)	χ^2 (df=1)p
Loss of pleasure	1301 (88%)	572 (98%)	χ^2=49,76 **p<.001**	558 (98%)	319 (99%)	χ2=1,33 p=.248
Diurnal mood variation	674 (47%)	301 (55%)	χ^2=10,49 **p<.001**	261 (47%)	157 (49%)	χ2=.263 p=.608
Suicidal ideation	1015 (70%)	503 (87%)	χ^2=67,57 **p<.001**	460 (81%)	284 (89%)	χ2=7,92 **p=.002**
Excessive self-reproach	1163 (81%)	548 (94%)	χ^2=56,76 **p<.001**	519 (95%)	294 (94%)	χ2=.23 p=.632
Poor concentration	1180 (82%)	550 (97%)	χ^2=72,74 **p<.001**	517 (95%)	315 (99%)	χ2=5,57 **p<.001**
Slowed activity	704 (52%)	370 (70%)	χ^2=47,16 **p<.001**	317 (64%)	197 (71%)	χ2=3,32 p=.069
Loss of energy	1257 (86%)	557 (96%)	χ^2=36,63 **p<.001**	552 (98%)	318 (97%)	χ2=.936 p=.333
Poor appetite	703 (50%)	362 (66%)	χ^2=39,87 **p<.001**	249 (47%)	208 (66%)	χ2=26,6 **p<.000**
Weight loss	370 (27%)	225 (44%)	χ^2=49,74 **p<.001**	132 (26%)	118 (40%)	χ2=19,19 **p<.000**
Increased appetite	331 (24%)	145 (27%)	χ2=2,04 p=.152	149 (29%)	88 (29%)	χ2=.007 P=.932
Weight gain	237 (18%)	113 (23%)	χ2=5,07 **p=.024**	118 (23%)	69 (23%)	χ2=.002 p=.967
Initial insomnia	484 (35%)	317 (60%)	χ2=90,01 **p<.001**	213 (41%)	197 (64%)	χ2=38,42 **p<.000**
Middle insomnia	386 (29%)	266 (53%)	χ2=88,72 **p<.001**	150 (30%)	164 (55%)	χ2=40,04 **p<.000**
Early morning waking	359 (27%)	255 (49%)	χ2=84,04 **p<.001**	161 (32%)	137 (47%)	χ2=18,80 **p=.016**
Excessive sleep	653 (49%)	236 (45%)	χ2=1,70 p=.192	275 (54%)	141 (47%)	χ2=3,60 p=.058
Diminished libido	783 (60%)	406 (82%)	χ2=75,44 **p<.001**	405 (84%)	238 (85%)	χ2=.353 p=.552

Total vary due to unknown/missing data

All the BD-I episode features that were significantly higher in the agitated group (p ≤.05) were included in the logistic regression analysis. The episode features that best predicted the presence of agitation in BD-I group were poor concentration (OR 2.694, 95% C.I: 1.176 – 6.172, p=.019), suicidal ideation (OR 1,809, 95% C.I: 1.078 – 2.710,

p=.023), middle insomnia (OR 1.820, 95% C.I: 1.277 – 2.593, p=.001), diminished libido (OR 1.939, 95% C.I: 1.361 – 2.764, p<.001), weight loss (OR 1.637, 95% C.I: 1.099- 2.438, p=.015) and slowed activity (OR 1.597, 95% C.I: 1.129 – 2.259, p=.008). This model considers 22% of the variance and correctly classified 75% of participants who have agitated features or not.

As for BD-I group, all the episode features with the ratios significantly higher (P ≤.05) in BD-II people with agitated features were included into a logistic regression analysis. The episode features that best predicted the presence of agitation were suicidal ideation (OR 1.711, 95% C.I: 1.011 – 2.897, p=.045), middle Insomnia (OR 2.234, 95% C.I: 1.499 – 3.330, p<.001) and poor appetite (OR 1.619, 95% C.I: 1.019- .2.572, p=.041). This model considers 14,5% of the variance and correctly classified 66,4% of participants who have agitated features or not.

4. Discussion

This study is carried out in a large clinical sample of UK, i.e. N=2123 BD-I, N=915 BD-II and N=712 MDD patients. Unlike the majority of the other studies that investigated agitated depression, we consider all the three mood disorders (MDD, BD-I and BD-II) in evaluating the rates of patients with agitated features in the worst depressive episode. We found AD episodes in the 29,3% of the whole sample. Specifically our study showed a prevalence of AD in Bipolar patients (BD-II 37%, BD-I 29% vs 21% MDD). This finding is in agreement with previous studies that compared the presence of agitated current episode both in the three disorders (Koukopulus et al., 2007; McIntyre et al., 2015; etc.) and in MDD and BD-II (Benazzi et al., 2002; Benazzi, 2004a; Dunner, & Tay, 1993; Takeshima, & Oka, 2013). However, regarding the proportions, our data are different from all the other studies. Moreover, our finding of the higher proportion of agitated depression in BD-II (37%) is inconsistent with some

previous researches in which the proportion was higher in BD-I (Koukopulus et al., 2007) or equal in BD-I and BD-II (McIntyre et al., 2015).

This heterogeneity of the results again highlights the problem in defining and evaluating agitated features of depression in different way. In particular, we used the definition of OPCRIT item of agitated activity (McGuffin et al., 1991) where the symptoms are easily recognizable by both the clinician and the patient (Olgiati et al., 2006). We suggest that the proportion of bipolar II in some other studies was lower (or equal) of bipolar I because some hypomanic symptoms of the AD (i.e. racing or crowded thoughts, irritability, talkativeness, dramatic descriptions of suffering, mood lability, etc.) may have been underestimated in BD-II as less severe (Judd et al., 2012).

Moreover, we investigated the possible correlations of AD in BD-I and II with demographic, lifetime and episode features. We don't consider MDD in this analysis to avoid confusion in the results. Indeed, many researches, which are already focused on AD in MDD, highlighted the association between psychomotor agitation and risk for mood-switching, suggesting that AD patients are often reclassified in the bipolar spectrum (Akiskal et al., 2005; Ansgst et al 2009; Benazzi et al., 2004; Benazzi, 2006; Biondi et al, 2005; Cassano et al., 2012; Fiedorowicz et al., 2011; Iwanami et al., 2015; Oligiati et al., 2006; Swann, 2013).

Some previous studies have suggested the relation between bipolar depression, panic comorbidity and more severity of the episode symptoms, in particular suicidal behaviour (Coryell, Endicott, Andreasen, Keller, & Clayton 1988; Dilsaver et al., 1997; Goodwin & Hoven, 2002; Kilbane, Gokbayrak, Galynker, Cohen, & Tross, 2009). In our finding the regression model in BD-I and II highlight the strong connection between AD episode and lifetime history of panic attacks. The combination agitated depression-panic attack makes the illness more severe. Moreover, patients suffering from panic disorder are usually treated with antidepressants (Furukawa, Watanabe, & Churchill

2006; Tsuboi & Masuko, 2001), which could result in the worsened of agitated symptomatology and the disorder's course (Koukopoulos, & Koukopoulos, 1999; Vázquez et al., 2013) and increase risk of suicide (Baldessarini et al., 2006a; Koukopulus & Koukopulus, 1999; 2007). Thus, the use of antidepressant in panic disorder should be closely monitored in order to individuate signs of excitatory phenomena and reconsider the treatment. On the other hand, patients with MDE, which have suffered from panic disorder, should be considered a high risk of agitated depression, and the risk/benefit of the administration of antidepressants carefully examined according to the psychiatric history of the individuals.

Moreover in BD-I, the result of regression model underlines the relationship between AD and the higher rate of people with depression as polarity of the first affective episode (81% vs 66%). Previous studies correlated the depressive polarity of the first onset with a lifetime depressive illness (i.e. more depressive episode in life), an increased risk of suicide attempts (Forty et al., 2009; Perugi et al., 2000) and DMI course (depression-mania-interval) with more severe symptomatology resistant to mood stabilizing treatments (Koukopoulos et al., 2013; McIntyre et al., 2014). In keeping with these studies, our finding highlights that when the onset of the bipolar illness is a depressive episode the course is more severe with high risk of agitated depression, and suicidal behaviour. Indeed, the DMI course has been associate with more resistance to mood stabilizer treatments.

Furthermore in BD-I, the regression model showed the link between agitated depression episode and lifetime dysphoric mania episodes. Dysphoric mania is an episode of mania in which the predominant mood is characterized by unease or mental discomfort (including low mood) that is consistent with mixed states definition of Kraepelin (1913, 1920) as well as agitated depression (Faedda et al., 2015; Goodwin &, Jamison, 2007; Koukopoulos, & Sani, 2014). Indeed, AD it is characterized by depressive polarity but

also by excitatory phenomena, such as agitation, restlessness and repetitive activity (Kraepelin, 1913; McGuffin et al., 1991; Oligiati et al., 2006). According patients with dysphoric mania episodes will most likely have agitated episode of depression, therefore much more mixed states in a lifetime (Swann et al., 1993). Thus, although depressive mood in mixed state can led the clinicians to administered antidepressants, these drugs should be avoid and the treatment with mood stabilizers or antipsychotics appears the most appropriate both for the agitated episode and for the course of the illness.

Previous studies underlined that patients with agitated depression were significantly more likely to have suicidal ideation and suicide attempts during the episode (Akiskal et al., 2005; Angst, Angst, & Stassen 1999; Balazs et al., 2006; Busch, Fawcett, & Jacobs 2003; Maris, 1985; Rihmer, Gonda, Balazs, & Faludi, 2008). Our results provide a further strong evidence that both in BD-I and -II, the "suicidal ideation" is more frequent in the AD compared with NON-AD episodes (BD-I 87% vs. 69% and BD-II 89% vs. 82%) in agreement with previous studies. Moreover, our results remark that patients, which experience agitated depression, have a higher risk of suicide attempts in lifetime (BD-I 59% vs. 44%, BD-II 56% vs. 47%).

The regression models showed that agitated features are closely associated with symptoms of biological nature - somatic symptoms - such as insomnia, diminished libido, weight loss, poor appetite and slowed activity in Bipolar Disorder, emphasizing the somatic features of agitated depression. This finding suggest that agitated and somatic symptoms in depression could share similar biological mechanisms. Therefore, deepening the aetiology of all the peculiar features associated with agitated symptoms episode may be helpful in recognizing and treating AD.

In addition, we found a higher proportion of alcohol misuse in agitated depression compared with non-AD. We suggest, in line with other studies (Himmelhoch, Mulla,

Neil, Detre, & Kupfer 1976; Sato et al., 2003), that patients with agitated features, having a more painful symptomatology, use alcohol to relieve symptoms (i.e. anguish, panic, insomnia, etc.) more often than non-AD patients. We also found a significant younger age at illness onset related to agitated features in BD-I (22 vs 24 years) in agreement with other studies (Koukopoulos, & Sani, 2014; Maj et al., 2003; Swann; 2013). Moreover, the agitated group had a significant higher rate of rapid cycling course both in BD-I and in BD-II (37% vs. 28%, and 54% vs 44%); this last finding is in agreement with some studies (Koukopoulos, & Sani, 2014; Maj et al., 2003) that correlated agitated depression with more affective recurrences and treatment resistant illness (Koukopulus et al., 2013).

Finally, in lifetime, patients with AD had a significant higher use of pharmacological drugs, suggesting once again the resistance to conventional treatments.

The potential limitation of this research is the difficulty in the conceptualization of the term "agitated depression". In fact, it is known that in the literature and in the different versions of the DSM this term has been the subject of controversy (Akiskal & Benazzi, 2004; Benazzi, 2004, Benazzi et al., 2004; Faedda et al., 2015; Koukopulus & Sani 2014; Maj et al.,2003; Olgiati et al.,2006). There is not a univocal definition of AD, and therefore there is not a univocal assessment. For this reason, our study that uses the OPCRIT criteria (McGuffin et al., 1991; Oligiati et al., 2006) could not include patients with other manic symptoms (i.e. talkativeness, or racing crowed thought). Nevertheless, OPCRIT definition (McGuffin et al., 1991) is consistent with original Kraepelin definition of AD in which mood and ideation are in the negative polarity and activity in a positive polarity (Akiskal & Benazzi, 2004). Indeed, Kraepelin distinguished agitated depression by depression with the flight of ideas characterized by mood and activity in negative polarity and ideation in positive polarity. Instead, Koukopulus includes in the

definition of mixed states (also called "agitated depression") both of the two forms of depression. This could explain the higher percentages of AD identified by Koukopulus in his studies (Koukopulus et al., 2007). Anyway, both agitated depression and depression with flight of ideas have common symptoms of manic polarity; for this reason, treatment should take into account the presence of excitatory symptoms as well as mixed states (Balazs et al., 2006; Baldessarini et al., 2006b; Baldessarini et al., 2013; Benazzi et al., 2004; Cuomo, Nikolova, Yalin, Arnone, Fagiolini, & Young 2017; Goldberg et al., 2007).

Our findings have a meaningful clinical relevance for the diagnosis and treatment of AD, a very severe form of depression occurring in a high percentage of patients with mood disorders and in particular in Bipolar Disorder. Moreover, AD is correlated with other excitatory phenomena (dysphoric mania, mixed state and panic attack), DMI course and suicidality. The occurrence of suicidal ideation along with agitation is a serious risk, because it might increase the probability of suicide in depression (Ağargün, Kara, & Solmaz, 1997; Angst et al., 1999; Busch et al., 2003). In agreement with several other studies (Forty et al., 2008; Godwin & Jamison, 2007; Sachs et al., 2007), our finding suggest again caution in the pharmacological treatment of bipolar depression; the symptomatology of any MDE, should be carefully evaluated in order to individuate agitation symptoms and provide an appropriate treatment. According to many authors, the most appropriate treatment of agitated depression could be the administration of anti-manic/mood stabilizer or antipsychotic drugs and to avoid antidepressants (Baldessarini et al., 2006b; Goodwin & Jamison, 2007; Koukopulus et al., 2007; Maj et al., 2003).

This research was intentionally entered at the end of the thesis project. It refers to the work done during the research period abroad; In particular, in collaboration with the

Bipolar Disorder Research Network (BDRN) of the University of Cardiff and Worcester, which allowed to work on a large clinical sample. The authors who contributed in a similar version of this work for the submission, are: Serra F [1], Knott S [2], Perry A [2], Forty L [3], Jones I [3], Craddock N [3], Gordon-Smith K [2], & Jones L [2]. (1= University of Padova; 2= University of Worcester; 3= University of Cardiff).

The methodological and clinical implication of this work are strictly related with the research done with University of Padua in these three years. Indeed, this last work emphasizes the need of appropriate tools for the correct diagnosis of each form of depression in order to avoid wrong treatment followed by dangerous consequences for the outcome of the illness and for the pain of the patients.

CHAPTER 8

OVERALL DISCUSSION

Mood disorders are among the most prevalent of all mental health diagnoses and their incidence has increased in last decades, becoming the most significant public health problem. Suicide is the most tragic consequence of mood disorders. It is associated with a mood disorder in 90% of cases and with a standard mortality ratio compared to the general population of 20:1.

A missed diagnosis of the specific episode means that the patient does not receive specific treatment, with dangerous consequences. Despite different efforts, under-diagnosis and under-treatment of mood disorders remain two serious problems worldwide.

Psychologists as well as Psychiatrists should develop the competence to detect the entire spectrum of mood disorders and should have the availability of appropriate tools for a differential diagnosis of every single case. The result of a good assessment is the possibility of treating the individual in an effective way.

The assessment is a wide spectrum evaluation the psychologists carry out in view of a proper treatment. Clinical interview, semi-structured interview and observation provide a large amount of information (i.e., exhaustivity), following an adaptive logic, and they take advantage from multiple channels (verbal, non-verbal). Despite the multiple pros of these tools, the problem of the inferences of the clinician can cause errors in later evaluation and diagnosis. Indeed, the clinician's evaluation could be affected by underestimation or overestimation of patient's symptoms. Regarding the psychometric approach, typically, the self-report questionnaires return a total score, which determines the impairment level of the individual. Psychometric testing allows the collection of a

lot of information in a short time; nevertheless, it is primarily data oriented, and the product is only a series of numeric scores that do not allow differentiating the symptomatology.

The CBA 2.0 suggested a valid option for a comprehensive approach to the psycho-diagnostic examination. Tests are used in a hierarchical way: in the initial stages, they explore several potentially problematic areas; then they investigate specific constructs. The CBA 2.0 provides a descriptive computerized report of the patient score which includes both the analysis of the validity and reliability of the test with high scores obtained in problematic areas and the positive replies to critical items. This attempt represented a great improvement in circumscribing problems to successive phases and proposing hypotheses concerning therapy. Despite this, it does not include the natural and logical flow of question-answer and it provides only quantitative information related to the unsatisfactory use of the resulting scores. Although there have been several attempts to apply adaptive clinical assessment, as far as we know, no system was able to combine adaptability, quantitative and qualitative information, and estimate error parameters through a probabilistic model.

A new methodological approach, called Formal Psychological Assessment (FPA), was designed to cope these problems.

The FPA was born in the University of Padova, and developed in an original way from the conjunction of two mathematical psychology theories: Knowledge Space Theory and Formal Concept Analysis.

FPA application to mood disorders assessment has been the core of this work, as well as a further step to overcome some obstacles encountered in the differential diagnosis of depression. Indeed, FPA could represent an important approach for improving case conceptualization and treatment implementation.

FPA allows for developing an instrument with multiples benefits based on the formal representation of the relationship between the "items" of a questionnaire and a given set of "clinical criteria".

FPA is potentially capable of maximizing the advantages of both semi-structured interviews and self-report questionnaires managing the problems of traditional assessment. The first step of FPA is the deterministic model, which consists of the construction of the Boolean matrix assigning to each item of the self-report scale, the subset of symptoms it investigates. The second step concerns the construction of the clinical structure from the attributes (symptoms) assignment, where each node represents a clinical state and its set of attributes. The clinical structure is a deterministic representation of the prerequisite relation among the items of the domain. However, a completely deterministic approach is inadequate for assessment in clinical practice since not all clinical states have the same probability and possible patient's answering errors may prevent a perfect correspondence between the observed response pattern and the actual clinical condition of the patient. Therefore, a probabilistic approach (i.e. the BLIM) was considered. It takes into account the probability of each clinical state and the probability of the false negative and the false positive rates for each item. The clinical structure, by means of the probabilistic weights obtained through the application of the BLIM, could be used to implement an adaptive algorithm.

In the present work of thesis, Formal Psychological Assessment has been applied both for the description of the self-report instruments used in clinical practice and for the construction of a new effective tool for Major Depressive Episode Assessment.

The first research, described in Chapter 4, aimed to define a practical application of FPA to illustrate procedural issues, discusses the advantages of the approach, and shows its potential for psychological assessment, relating to depression. Specifically, in this first part FPA was applied to analyse the "item content" of the most used self-report

questionnaires for the depression's evaluation. Indeed FPA allowed creating relations between "item content" and "diagnostic criteria" to the assessment of Major Depressive Episode (MDE). In keeping with previous researches, the main task of the first study was to underline the strengths and the weaknesses of widely used depression tools, in relation of their ability to investigate all the symptoms of MDE. To achieve this aim we used the main concept of FPA described in the third Chapter. The clinical criteria used came from the DSM-5, the literature and the Beck and Seligman's theories of depression. Through FPA, we highlighted each self-report questionnaire's strengths and weaknesses in terms of correspondence to a set of diagnostic and clinical criteria. None of these tools was able to investigate alone the whole set of symptoms considered essential for the evaluation of MDE by the experts in the field of mood disorders. This methodology allowed to eliminate useless redundancy and to increase efficiency. The careful analysis of the items of the questionnaires has allowed creating the skeleton for the construction of a new instrument for the major depressive episode.

Flexibility is another crucial advantage of FPA: the set of attributes (symptoms or clinical criteria) could be easily modified or updated according to new versions of DSM or to different theoretical approaches, while the methodology remains equally effective and reliable.

The second research, presented in Chapter 5, aimed to construct a quantitative-qualitative tool that investigates all Major Depressive Episode (MDE) clinical features and provides qualitative information (and not just a score) to differentiate patients with the same score but different symptoms as well as different severity of psychopathology. In this research, we applied the FPA framework and fruitfully used the main concepts described in the third Chapter. The main purpose of this study was to create an innovative tool that takes into account the strengths and weaknesses of the tools analysed in the previous research. The research explains the construction of 41

dichotomous items based on 23 clinical criteria of major depressive episodes from the DSM-5 and the literature. The Quantitative and Qualitative Evaluation of Depressive Symptomatology (QuEDS) was tested and validated in both clinical and non-clinical sample. The QuEDS takes into account all of the positive responses of the subject, which are closely linked to the symptoms through the FPA (MDE clinical criteria). In this way, clinicians will no longer be bound to the patient's score but will be interested in the patient's clinical state that is the set of items to which the patient responded positively, along with the set of symptoms investigated by those items. Such information is already present in the items, but it is hidden by a classical testing methodology that considers the questionnaire score to be the most relevant output used by the clinicians. The qualitative differences in symptoms between the two patients are highly relevant for a correct diagnosis and for future psychological and pharmacological treatment.

It is noteworthy to observe that physicians sometimes prescribe antidepressants without carefully analysing the individual's depressive symptoms, also when these drugs can be very dangerous and can increase the risk of suicide in patients with mixed depression.

Therefore, the output of the proposed tool (QuEDS) is the patient's clinical state; it is no longer the score. From a clinical point of view, a qualitative self-report tool overcomes the cut-off limit, which can be helpful just to have an idea about the test score, but it cannot be mistaken for a correct estimate of a person's symptomatology.

Thus, the QuEDS could be a useful contribution for many reasons: first, for the broad spectrum of clinical criteria investigated by the test. Second, for the importance given to qualitative information about the symptoms (through the patient's clinical state) and not only to the score. Third, for the relevance given to the differences in symptoms and especially in their severity. Fourth for its appreciable validity and reliability results. Finally, for the application of an adaptive logic.

Indeed, the third research, presented in the Chapter 6, showed the implementation of the three algorithms on the three sub-dimensions of the QuEDS. We derived the clinical structures of the three sub-matrices related to the three sub-factors cognitive, somatic, and affective (Chapter 5, Table 5.6) of depression. The three models were tested on the data of all the participants who completed the written form of the QuEDS in order to estimate the parameters (the false negative and the false positive rates for each item) and the fit indexes. The following step involved the simulation and aimed to reproduce the writing version of QuEDS in an adaptive way. The ATS-PD algorithm took as input the clinical structure of each sub-scale of QuEDS (cognitive, somatic, affective) and the parameters estimates (β, μ), and reproduced the assessment, evaluating whether the adaptive form of QuEDS could generate the same response pattern of a subject who answered the written version (system efficacy), but with a smaller number of questions (system efficiency). The ATS-PD implementation showed that, through a computerized system of FPA, the tool allows to administer a smaller number of items to the individual without loss of measurement precision and according with his previous answer. In this way, a sequence of question is asked and at the end, one of the all clinical state of the structure should achieve a high probability value (with a fixed cut-off .70). This clinical state represents the most likely symptomatic representation of the patient's situation regarding a specific sub-factor. At this point, the algorithm stops and provides the clinician with the score, the response pattern, and the attributes configuration (all the symptoms complained by patients). The adaptive version of QuEDS mimics the semi-structured interviews process that allowed examining in depth only the individual's symptomatic areas.

The QuEDS allows for an adaptive, quantitative and qualitative evaluation of depressive symptomatology. Adaptive because, based on the structure and the algorithm, it selects each question to maximize the collectable information. Quantitative because it could

provide a numerical score with the level of impairment. Qualitative because it provides information about all the subjects' symptoms allowing the differentiation among different types of Major Depressive Episode.

In evaluating depression, an adaptive tool able to going beyond the numerical score is essential. In fact, clinicians have to face different depressive symptoms that often require a different diagnosis and appropriate specific treatment. The case of agitated depression, or mixed depression, represents perhaps the most important. In fact, agitated depression is classified as a mixed episode, with both depressive and manic symptoms, and for this reason it cannot be treated in the same way as the typical depressive episode.

The fourth research, in the Chapter 7, underlined the clinical relevance of Agitated Depression (AD) in mood disorders. It appears to be a form of depression with very severe symptomatology (i.e. associated with more depressive episodic symptoms); in particular somatic symptoms (i.e. insomnia, poor appetite, etc.) and suicide ideation. Agitated patients have a significant higher use of psychotropic drugs in lifetime and higher rate of panic comorbidity and suicidality compared with depressive patients without agitation. In agreement with several other studies, our finding of this last research suggest again caution in the pharmacological treatment of bipolar depression. The recognition and the differential diagnosis of AD is crucial to avoid erroneous and dangerous pharmacological treatments. Indeed, antidepressant administration could worsen the excitatory symptoms resulting in the failure to relieve the patient's pain and, as more serious consequence, they could increase the risk of suicide. According to many authors, the most appropriate treatment of agitated depression could be the administration of anti-manic/mood stabilizer or antipsychotic drugs.

The methodological and clinical implication of this last work, carried out in England (with the University of Cardiff and Worcester), are strictly related with the research

done with University of Padua in these three years. Indeed, it emphasizes the need of the correct diagnosis of agitated depression in order to avoid wrong treatment followed by dangerous consequences for the outcome of the illness and for patients' pain. To achieve this goal we need to develop new effective tools that are able to differentiate major depressive episode with mixed features from major depressive episode without mixed features. The QuEDS could represent a valid support in the assessment phase since the recognition of agitated depression is not initially easy; indeed, both the patient and the clinician often underestimate the excitatory symptoms (such as irritability, agitation, mood lability, anguish, racing thought) considering only depressive symptoms. Notwithstanding the various pros of the self-report tools, it is important to note that for a complete assessment the importance of the interviews cannot be ignored for the ability to investigate non-verbal behavior.

Future implementations will be oriented to the improvement of the new created tool, considering all the possible symptoms of mixed depression and updating the adaptive form of the QuEDS. Indeed, the quality of clinical evaluation is crucial for both diagnosis and treatment and, as stated in this work, an erroneous psycho-diagnostic evaluation could result in misdiagnosis and therefore in therapeutic failures.

REFERENCES

Abela, B. L. (2008). Handbook of Depression in Children and Adolescents. *Psychological Medicine*, 38(12), 1816- 1817.

Abramson, L. Y., Seligman, M. E., & Teasdale, J. D. (1978). Learned helplessness in humans: critique and reformulation. *Journal Abnormal Psychology*, 87(1), 49-74.

Ağargün, M. Y., Kara, H., & Solmaz, M. (1997). Sleep disturbances and suicidal behaviour in patients with major depression. *The Journal of clinical psychiatry*, 58(6), 249-251.

Akiskal, H. S., & Benazzi, F. (2004). Validating Kraepelin's two types of depressive mixed states: "depression with flight of ideas" and "excited depression". *World Journal of Biological Psychiatry*, 5:107–113.

Akiskal, H. S., & Benazzi, F. (2003). Family history validation of the bipolar nature of depressive mixed states. *Journal of affective disorders*, 73(1), 113-122.

Akiskal, H. S., Cassano, G. B. (1997). *Dysthymia and the Spectrum of Chronic Depressions*. Guilford Press, New York.

Akiskal, H. S., & Pinto, O. (1999). The evolving bipolar spectrum: prototypes I, II, III, and IV. *Psychiatric Clinics of North America*, 22(3), 517-534.

Akiskal, H. S., Benazzi, F., Perugi, G., & Rihmer, Z. (2005). Agitated "unipolar" depression re-conceptualized as a depressive mixed state: implications for the antidepressant-suicide controversy. *Journal of affective disorders*, 85(3), 245-258.

Al Busaidi, Z. Q. (2010). The concept of somatisation: a cross-cultural perspective. *Sultan Qaboos University Medical Journal*, 10(2), 180.

Alexopoulos, G. S., Hoptman, M. J., Yuen, G., Kanellopoulos, D., Seirup, J. K., Lim, K. O., & Gunning, F. M. (2013). Functional connectivity in apathy of late-life depression: a preliminary study. *Journal of affective disorders*, 149(1), 398-405.

Alloy, L. B., Peterson, C., Abramson, L. Y., & Seligman, M. E. (1984). Attributional style and the generality of learned helplessness. *Journal of personality and social psychology*, 46(3), 681.

American Psychiatric Association [APA]. (1995). DSM-IV-TR. *Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition, Text Revision ed.)*. Washington D. C.: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Andreasen, N. C., Grove, W. M. (1982). The classification of depression: Traditional views versus mathematical approaches. *The American journal of psychiatry*. 139, 45–52

Angst, J., Gamma, A., Benazzi, F., Ajdacic, V., & Rössler, W. (2009). Does psychomotor agitation in major depressive episodes indicate bipolarity? *European archives of psychiatry and clinical neuroscience*, 259(1), 55-63.

Angst, J. (1998). The emerging epidemiology of hypomania and bipolar II disorder. *Journal of affective disorders*, 50(2), 143-151.

Angst, J., Angst, F., & Stassen, H. H. (1999). Suicide risk in patients with major depressive disorder. *Journal of clinical psychiatry*. 60 (Suppl. 2), 57–62

Annen, S., Roser, P., & Brüne, M. (2012). Nonverbal behavior during clinical interviews: similarities and dissimilarities among schizophrenia, mania, and depression. The Journal of nervous and mental disease, 200(1), 26-32.

Augustin, T., Hockemeyer, C., Kickmeier-Rust, M. D., Podbregar, P., Suck, R., & Albert, D. (2013). The simplified updating rule in the formalization of digital educational games. *Journal of Computational Science*, 4, 293-303.

Baek, S. G. (1997). Computerized adaptive testing using the partial credit model for attitude measurement. *Objective measurement: Theory into practice*, 4, 37-43

Balázs, J., Benazzi, F., Rihmer, Z., Rihmer, A., Akiskal, K. K., & Akiskal, H. S. (2006). The close link between suicide attempts and mixed (bipolar) depression: implications for suicide prevention. *Journal of affective disorders*, 91(2), 133-138. doi: 10.1016/j.jad.2005.12.049

Baldessarini, R. J., Faedda, G. L., Offidani, E., Vázquez, G. H., Marangoni, C., Serra, G., & Tondo, L. (2013). Antidepressant-associated mood-switching and transition from

unipolar major depression to bipolar disorder: a review. *Journal of affective disorders*, 148(1), 129-135.

Baldessarini, R. J., Pompili, M., & Tondo L. (2006a). Suicide in bipolar disorder: Risks and management. *CNS Spectrums*, 11(6): 465-471.

Baldessarini, R. J., Tondo, L., Davis, P., Pompili, M., Goodwin, F. K., and Hennen, J. (2006b). Decreased risk of suicides and attempts during long-term lithium treatment: a meta-analytic review. *Bipolar Disorder*, 8(5 Pt 2), 625–639. doi: 10.1111/j.1399-5618.2006.00344.x

Baldessarini, R. J., Vieta, E., Calabrese, J. R., Tohen, M., & Bowden, C. L. (2010). Bipolar depression: overview and commentary. *Harvard review of psychiatry*, 18(3), 143-157.

Balsamo, M., & Saggino, A. (2007). Test per l'assessment della depressione nel contesto italiano: un'analisi critica. *Psicoterapia Cognitivo Comportamentale*, 13(2), 167-199.

Barlow, D. H. (2004). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. Guilford press.

Batterham, P. J., Ftanou, M., Pirkis, J., Brewer, J. L., Mackinnon, A. J., Beautrais, A., & Christensen, H. (2015). A systematic review and evaluation of measures for suicidal ideation and behaviors in population-based research. *Psychological Assessment*, 27, 501–512. doi: 10.1037/pas0000053

Bauer, M., & Pfennig, A. (2005). Epidemiology of bipolar disorders. *Epilepsia*, 46(s4), 8-13.

Bayram, N., & Bilgel, N. (2008). The prevalence and socio-demographic correlations of depression, anxiety and stress among a group of university students. *Social psychiatry and psychiatric epidemiology*, *43*(8), 667-672.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory-II*. San Antonio, 78(2), 490-8.

Beck, A. T. (1991). Cognitive therapy. A 30-year retrospective. *American Psychologist*, 46(4): 368-375.

Beck, A. T. (2005). The current state of cognitive therapy: a 40-year retrospective. *Archives of General Psychiatry*, 62(9): 953-959.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical psychology review*, 8(1), 77-100.

Beigel, A., & Murphy, D. L. (1971). Unipolar and bipolar affective illness: Differences in clinical characteristics accompanying depression. *Archives of General Psychiatry*, 24(3), 215-220.

Benazzi, F. (2006). A continuity between bipolar II depression and major depressive disorder? *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 30(6), 1043-1050.

Benazzi, F. (2004a). Agitated depression: a valid depression subtype? *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 28(8), 1279-1285.

Benazzi, F. (2004b). Is depressive mixed state a transition between depression and hypomania? *European archives of psychiatry and clinical neuroscience*, 254(2), 69-75.

Benazzi, F. (2007). Bipolar disorder—focus on bipolar II disorder and mixed depression. *The Lancet*, 369(9565), 935-945.

Benazzi, F., & Akiskal, H. (2005). Irritable-hostile depression: further validation as a bipolar depressive mixed state. *Journal of affective disorders*, 84(2), 197-207.

Benazzi, F., Helmi, S., & Bland, L. (2002). Agitated depression: unipolar? Bipolar? Or both? *Annals of Clinical Psychiatry*, 14(2), 97-104.

Benazzi, F., Koukopoulos, A., & Akiskal, H. S. (2004). Toward a validation of a new definition of agitated depression as a bipolar mixed state (mixed depression). *European Psychiatry*, 19(2), 85-90.

Benca, R. M., & Peterson, M. J. (2008). Insomnia and depression. *Sleep medicine*, 9, S3-S9.

Bennabi, D., Vandel, P., Papaxanthis, C., Pozzo, T., & Haffen, E. (2013). Psychomotor retardation in depression: a systematic review of diagnostic, pathophysiologic, and therapeutic implications. *BioMed research international*.

Berrios, R., Kellett, S., Fiorani, C., & Poggioli, M. (2016). Assessment of identity disturbance: Factor structure and validation of the Personality Structure Questionnaire in an Italian sample. *Psychological assessment*, 28(4), e27.

Bertolotti, G., Zotti, A. M., Michielin, P., Vidotto, G., & Sanavio, E. (1990).A computerized approach to cognitive behavioural assessment: An introductionto CBA-2.0 primary scales. *Journal of Behavior Therapy and Experimental Psychiatry*, 21(1), 21-27.

Bidzi, E. J. (1984). Stress factors in affective diseases. *The British Journal of Psychiatry*, 144(2), 161-166.

Bijl, R. V., Ravelli, A., & Van Zessen, G. (1998). Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social psychiatry and psychiatric epidemiology*, 33(12), 587-595.

Bijl, R. V., & Ravelli, A. (2000). Current and residual functional disability associated with psychopathology: findings from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Psychological medicine*, 30(03), 657-668.

Billiard, M., Dolenc, L., Aldaz, C., Ondze, B., & Besset, A. (1994). Hypersomnia associated with mood disorders: a new perspective. *Journal of Psychosomatic Research*, 38, 41-47.

Biondi, M., Picardi, A., Pasquini, M., Gaetano, P., & Pancheri, P. (2005). Dimensional psychopathology of depression: detection of an 'activation'dimension in unipolar depressed outpatients. *Journal of affective disorders*, 84(2), 133-139.

Birkhoff, G. (1937). Rings of sets. *Duke Mathematical Journal*, 3, 443–454.

Birkhoff, G. (1967). Lattice Theory. Providence, R.I.: American Mathematical Society Colloquium Pubblication. (No. XXV)

Bland, R. C. (1997). Epidemiology of affective disorders: a review. *The Canadian Journal of Psychiatry*, 42(4), 367-377.

Bocquier, A., Pambrun, E., Dumesnil, H., Villani, P., Verdoux, H., Verger, P. (2013) Physicians' characteristics associated with exploring suicide risk among patients with

depression: a French panel survey of general practitioners. *PLoS One*. doi: 10.1371/journal.pone.0080797.

Bokhari, E., & Hubert, L. (2015). A new condition for assessing the clinical efficiency of a diagnostic test. *Psychological assessment*, 27(3), 745.

Bottesi, G., Spoto, A., Freeston, M. H., Sanavio, E., & Vidotto, G. (2015a). Beyond the score: clinical evaluation through formal psychological assessment. *Journal of personality assessment*, 97(3), 252-260.

Bottesi, G., Ghisi, M., Altoè, G., Conforti, E., Melli, G., & Sica, C. (2015b). The Italian version of the Depression Anxiety Stress Scales-21: factor structure and psychometric properties on community and clinical samples. *Comprehensive psychiatry*, 60, 170-181.

Bracht, T., Federspiel, A., Schnell, S., Horn, H., Höfle, O., Wiest, R., & Walther, S. (2012). Cortico-cortical white matter motor pathway microstructure is related to psychomotor retardation in major depressive disorder. *PloS One*, 7(12), e52238.

Brent, D. A., & Mann, J. J. (2005, February). Family genetic studies, suicide, and suicidal behavior. In American Journal of Medical Genetics Part C: Seminars in Medical Genetics (Vol. 133, No. 1, pp. 13-24). Wiley Subscription Services, Inc., A Wiley Company.

Brockington, I. F., Helzer, J. E., Hillier, V. F., & Francis, A. F. (1982). Definitions of depression: concordance and prediction of outcome. *American Journal Psychiatry*, 139(8), 1022-1027.

Busch, K. A., Fawcett, J., & Jacobs, D. G. (2003). Clinical correlates of inpatient suicide. *The Journal of clinical psychiatry*.

Calabrese, J. R., Rapport, D. J., Findling, R. L., Shelton, M. D., & Kimmel, S. E. (2000). Rapid-cycling bipolar disorder. In Bipolar disorders (pp. 89-90). Springer Netherlands.

Campo, J. V. (2012). Annual Research Review: Functional somatic symptoms and associated anxiety and depression–developmental psychopathology in pediatric practice. *Journal of Child Psychology and Psychiatry*, 53(5), 575-592.

Carroll, B. J., Feinberg, M., Smouse, P. E., Rawson, S. G., & Greden, J. F. (1981). The Carroll rating scale for depression. I. Development, reliability and validation. *The British Journal of Psychiatry*, 138(3), 194-200.

Cassano, G. B., & Castrogiovanni, P. (1982). SAD, Scala di autovalutazione della depressione. *Milan: International Committee for Prevention and Treatment of Depression. La Condizione Depressiva*, 483-486.

Cassano, G. B., Rucci, P., Benvenuti, A., Miniati, M., Calugi, S., Maggi, L., Pini, F., Kupfer, D. J., Fagiolini, A., & Frank, E. (2012). The role of psychomotor activation in discriminating unipolar from bipolar disorders: a classification-tree analysis. *The Journal of clinical psychiatry*, 73(1), 22-28.

Chen, Y. W., & Dilsaver, S. C. (1996). Lifetime rates of suicide attempts among subjects with bipolar and unipolar disorders relative to subjects with other Axis I disorders. *Biological psychiatry*, 39(10), 896-899.

Chen, C. H., Lee, C. S., Lee, M. T. M., Ouyang, W. C., Chen, C. C., Chong, M. Y., & Chiu, N. Y. (2014). Variant GADL1 and response to lithium therapy in bipolar I disorder. *New England Journal of Medicine*, 370(2), 119-128.

Clark, D. A., & Beck, A. T. (2010). Cognitive theory and therapy of anxiety and depression: convergence with neurobiological findings. *Trends in Cognitive Sciences*, 14 (9): 418-424.

Clark, D. M. (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *International review of psychiatry*, 23(4), 318-327.

Connolly, C. G., Ho, T. C., Blom, E. H., LeWinn, K. Z., Sacchet, M. D., Tymofiyeva, O., & Yang, T. T. (2017). Resting-state functional connectivity of the amygdala and longitudinal changes in depression severity in adolescent depression. *Journal of Affective Disorders*, 207, 86-94.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, 78(1), 98.

Coryell, W., & Tsuang, M. T. (1985). Major depression with mood-congruent or mood-incongruent psychotic features: outcome after 40 years. *American Journal of Psychiatry*, 142(4), 479-482.

Coryell, W., Endicott, J., Andreasen, N. C., Keller, M. B., & Clayton, P. J. (1988). Depression and panic attacks: the significance of overlap as reflected in follow-up and family study data. *The American journal of psychiatry*, 145(3), 293.

Coryell, W., Keller, M., Endicott, J., Andreasen, N., Clayton, P., & Hirschfeld, R. (1989). Bipolar II illness: course and outcome over a five-year period. *Psychological Medicine*, 19(01), 129-141.

Cross-National Collaborative Group. (1992) The changing rate of major depression: cross-national comparisons. JAMA 1992; 268:3098-3105.

Cuomo, A., Nikolova, V. L., Yalin, N., Arnone, D., Fagiolini, A., & Young, A. H. (2017). Pharmacological treatment of mixed states. *CNS spectrums*, 22(2), 186-195.

De Dios, C., Goikolea, J. M., Colom, F., Moreno, C., & Vieta, E. (2014). Bipolar disorders in the new DSM-5 and ICD-11 classifications. *Revista de Psiquiatría y Salud Mental* (English Edition), 7(4), 179-185.

Demontis, F., Serra, F., & Serra, G. (2017). Antidepressant-induced Dopamine Receptor Dysregulation: A Valid Animal Model of Manic-Depressive Illness. *Current neuropharmacology*, 15(3), 417-423.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1-38.

Devanand, D. P. (2014). Dysthymic disorder in the elderly population. *International psychogeriatrics*, 26(01), 39-48.

Diener, E., & Fujita, F. (1995). Resources, personal strivings, and subjective well-being: a nomothetic and idiographic approach. *Journal of personality and social psychology*, 68(5), 926.

Dilsaver, S. C., Chen, Y. W., Swann, A. C., Shoaib, A. M., Tsai-Dilsaver, Y., & Krajewski, K. J. (1997). Suicidality, panic disorder and psychosis in bipolar depression, depressive-mania and pure-mania. *Psychiatry research*, 73(1), 47-56.

Dinger, U., Barrett, M. S., Zimmermann, J., Schauenburg, H., Wright, A. G., Renner, & Barber, J. P. (2015). Interpersonal problems, dependency, and self-criticism in major depressive disorder. *Journal of Clinical Psychology*, 71, 93–104. doi: 10.1002/jclp.22120

Doignon, J. P., & Falmagne, J. C. (1985). Spaces for the assessment of knowledge. *International journal of man-machine studies*, *23*(2), 175-196.

Doignon, J. P., & Falmagne, J. C. (1999). Knowledge Spaces. Berlin; Heidelberg: Springer-Verlag. doi: 10.1007/978-3-642-58625-5

Doignon, J. P., & Falmagne, J. C. (2011). *Learning spaces*.

Donadello, I., Spoto, A., Sambo, F., Badaloni, S., Granziol, U., & Vidotto, G. (2016). ATS-PD: an adaptive testing system for psychological disorders. *Educational and Psychological Measurement*, 0013164416652188.

Dumont, M., & Provost, M. A. (1999). Resilience in adolescents: Protective role of social support, coping strategies, self-esteem, and social activities on experience of stress and depression. *Journal of youth and adolescence*, 28(3), 343-363.

Dunner, D. L., & Tay, L. K. (1993). Diagnostic reliability of the history of hypomania in bipolar II patients and patients with major depression. *Comprehensive Psychiatry*, 34(5), 303-307.

Eaton WW, Muntaner C, Smith C, Tien A, Ybarra M. (2004) Center for Epidemiologic Studies Depression Scale: Review and revision (CESD and CESD-R). In: Maruish ME, editor. The Use of Psychological Testing for Treatment Planning and Outcomes Assessment.3rd ed. Mahwah, NJ: Lawrence Erlbaum. pp. 363–377.

Editorial Lancet. (2012) Depression and the global economic crisis: is there hope? *Lancet* 380 : 1203.

Eggen, T. J. H. M., & Straetmans, G. J. J. M. (2000). Computerized adaptive testing for classifying examinees into three categories. *Educational and Psychological measurement*, 60(5), 713-734.

Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire (junior and adult)*. Hodder and Stoughton.

Faedda, G. L., Marangoni, C., & Reginaldi, D. (2015). Depressive mixed states: A reappraisal of Koukopoulos′ criteria. *Journal of affective disorders*, 176, 18-23.

Falmagne, J. C., Koppen, M., Villano, M., Doignon, J. P., & Johannesen, L. (1990). Introduction to knowledge spaces: How to build, test, and search them. *Psychological Review*, 97(2), 201.

Falmagne, J. C., and Doignon, J. P. (2011). "Knowledge spaces," in Learning Spaces (Berlin; Heidelberg: Springer), 43–60. doi: 10.1007/978-3-642-01039-2_3

Faraone, S. V., Tsuang, M. T., & Tsuang, M. T. (1995). Methods in psychiatric genetics. *Textbook in psychiatric epidemiology*, 81-134

Faravelli, C., Albanesi, G., & Poli, E. (1986). Assessment of depression: a comparison of rating scales. *Journal of Affective Disorders*, 11, 245–253. doi: 10.1016/0165-0327(86)90076-5

Fava, M., Ball, S., Nelson, J. C., Sparks, J., Konechnik, T., Classi, P., Dube, S., & Thase, M. E. (2014). Clinical relevance of fatigue as a residual symptom in major depressive disorder. *Depression and anxiety*, 31(3), 250-257.

Fava, G. A., Ruini, C., & Rafanelli, C. (2004). Psychometric theory is an obstacle to the progress of clinical research. *Psychotherapy and psychosomatics*, 73(3), 145-148.

Fava, M., & Rosenbaum, J. F. (1999). Anger attacks in patients with depression. *The Journal of clinical psychiatry*.

Ferentinos, P., Paparrigopoulos, T., Rentzos, M., Zouvelou, V., Alexakis, T., & Evdokimidis, I. (2011). Prevalence of major depression in ALS: comparison of a semi-structured interview and four self-report measures. *Amyotrophic lateral sclerosis*, 12(4), 297-302.

Fiedorowicz, J. G., Endicott, J., Leon, A. C., Solomon, D. A., Keller, M. B., & Coryell, W. H. (2011). Subthreshold hypomanic symptoms in progression from unipolar major depression to bipolar disorder. *American Journal of Psychiatry*, 168(1), 40-48.

Finkelman, M. D., Smits, N., Kim, W., & Riley, B. (2012). Curtailment and stochastic curtailment to shorten the CES-D. *Applied Psychological Measurement*, 36, 632-658.

Fiquer, J. T., Boggio, P. S., & Gorenstein, C. (2013). Talking bodies: Nonverbal behavior in the assessment of depression severity. *Journal of affective disorders*, 150(3), 1114-1119.

First, M. B., Gibbon, M., Spitzer, R. L., & Williams, J. B. (1996). User's guide for the structured clinical interview for DSM-IV axis I Disorders—Research version. New York: Biometrics Research Department, New York State Psychiatric Institute.

First, M. B., Benjamin, L. S., Gibbon, M., Spitzer, R. L., & Williams, J. B. (1997). Structured clinical interview for DSM-IV Axis II personality disorders. American Psychiatric Press.

Fliege, H., Becker, J., Walter, O. B., Bjorner, J. B., Klapp, B. F., & Rose, M. (2005). Development of a computer-adaptive test for depression (D-CAT). *Quality of life Research*, 14(10), 2277-2291.

Forouzanfar, M. H., Alexander, L., Anderson, H. R., Bachman, V. F., Biryukov, S., Brauer, M., & Delwiche, K. (2015). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 386(10010), 2287-2323.

Forty, L., Jones, L., Jones, I., Smith, D. J., Caesar, S., Fraser, C., Gordon-Smith, K., Hyde, S., & Craddock, N. (2009). Polarity at illness onset in bipolar I disorder and clinical course of illness. *Bipolar disorders*, 11(1), 82-88.

Forty, L., Smith, D., Jones, L., Jones, I., Caesar, S., Cooper, C., Fraser, C., Gordon-Smith, K., Hyde, S., Farmer, A., McGuffin, P., & Craddock, N. (2008). Clinical differences between bipolar and unipolar depression. *The British Journal of Psychiatry*, 192(5), 388-389

Frank, E., & Thase, M. E. (1999). Natural history and preventative treatment of recurrent mood disorders. *Annual review of medicine*, 50(1), 453-468.

Fried, E. I., & Nesse, R. M. (2014). The impact of individual depressive symptoms on impairment of psychosocial functioning. *PLoS One*, 9(2), e90311.

Furukawa, T. A., Watanabe, N., & Churchill, R. (2006). Psychotherapy plus antidepressant for panic disorder with or without agoraphobia. *The British Journal of Psychiatry*, 188(4), 305-312.

Gaillard, R., Gourion, D., Llorca, P. M. (2013) Anhedonia in depression. *Encephale*. 39: 296–305. doi: 10.1016/j.encep.2013.07.001 PMID: 23937895

Ganter, B., & Wille, R. (1999). Formal Concept Analysis: Mathematical Foundations. Transl. by Cornelia Franzke from the German. Berlin; Heidelberg: Springer-Verlag. doi: 10.1007/978-3-642-59830-2

Ghaemi, S. N., Hsu, D. J., Soldani, F., & Goodwin, F. K. (2003). Antidepressants in bipolar disorder: the case for caution. *Bipolar disorders*, 5(6), 421-433.

Gibbons, R. D., Clark, D. C., Cavanaugh, S., and Davis, J. M. (1985). Application of modern psychometric theory in psychiatric research. *Journal of Psychiatry Research*. 19, 43–55. doi: 10.1016/0022-3956(85)90067-6

Gibbons, R. D., Weiss, D. J., Kupfer, D. J., Frank, E., Fagiolini, A., Grochocinski, V. J., & Immekus, J. C. (2008). Using computerized adaptive testing to reduce the burden of mental health assessment. *Psychiatric Services*, 59(4), 361-368.

Goldberg, J. F., Perlis, R. H., Ghaemi, S. N., Calabrese, J. R., Bowden, C. L., Wisniewski, S., & Thase, M. E. (2007). Adjunctive antidepressant use and symptomatic recovery among bipolar depressed patients with concomitant manic symptoms: findings from the STEP-BD. *American Journal of Psychiatry*, 164(9), 1348-1355.

Goldring, N., & Fieve, R. R. (1984). Attempted suicide in manic-depressive disorder. *American journal of psychotherapy*, 38(3), 373-383.

Goodman & Gilman's Pharmacological Basis of Therapeutics, 12th Edition (2011) by The McGraw-Hill Companies.

Goodwin, F. K., & Jamison, K. R. (2007). *Manic-depressive illness: bipolar disorders and recurrent depression* (Vol. 1). Oxford University Press.

Goodwin, R. D., & Hoven, C. W. (2002). Bipolar–panic comorbidity in the general population: prevalence and associated morbidity. *Journal of Affective Disorders*, 70(1), 27-33.

Grayce, C. J. (2013). A commercial implementation of knowledge space theory in college general chemistry. In Knowledge spaces (pp. 93-113). Springer Berlin Heidelberg.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24(2), 95-112.

Grossberg, J. M. (1964). Behavior therapy: a review. Psychological Bulletin 62, 73–88. doi: 10.1037/h0041033.

Groth-Marnat, G. (2009). *Handbook of Psychological Assessment.* Hoboken, NJ: John Wiley & Sons.

Gunnell, D., & Middleton, N. (2003). National suicide rates as an indicator of the effect of suicide on premature mortality. *The Lancet*, 362, pp.961-962.

Hamilton M. A. (1960). Rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23(1) 56–62. pmid:14399272.

Hamoen, A. B., Redlich, E. M., & Weerd, A. W. (2014). Effectiveness of cognitive behavioral therapy for insomnia: influence of slight-to-moderate depressive symptom severity and worrying. *Depression and anxiety*, 31(8), 662-668.

Hantouche, E. G., & Akiskal, H. S. (2005). Bipolar II vs. unipolar depression: psychopathologic differentiation by dimensional measures. *Journal of affective disorders*, 84(2), 127-132.

Harris, B., Young, J., & Hughes, B. (1984). Appetite and weight change in patients presenting with depressive illness. *Journal of affective disorders*, 6(3), 331-339.

Harris, E. C., & Barraclough, B. (1997). Suicide as an outcome for mental disorders. A meta-analysis. *The British Journal of Psychiatry*, 170(3), 205-228.

Hasin, D. S., Goodwin, R. D., Stinson, F. S., & Grant, B. F. (2005). Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. *Archives of general psychiatry*, 62(10), 1097-1106.

Hathaway, S. R., & McKinley, J. C. (1942). A multiphasic personality schedule (Minnesota): III. The measurement of symptomatic depression. *The Journal of Psychology*, *14*(1), 73-84.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist*, 42(11), 963-974.

Henderson, S. E., Johnson, A. R., Vallejo, A. I., Katz, L., Wong, E., & Gabbay, V. (2013). A Preliminary Study of White Matter in Adolescent Depression: Relationships with Illness Severity, Anhedonia, and Irritability. *Frontiers in Psychiatry*, 4, 152. http://doi.org/10.3389/fpsyt.2013.00152

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British journal of clinical psychology*, 44(2), 227-239.

Hewitt, P. L., & Flett, G. L. (1993). Dimensions of perfectionism, daily stress, and depression: a test of the specific vulnerability hypothesis. *Journal of Abnormal Psychology* 102:58. doi: 10.1037/0021-843X.102.1.58.

Himmelhoch, J. M., Mulla, D., Neil, J. F., Detre, T. P., & Kupfer, D. J. (1976). Incidence and significance of mixed affective states in a bipolar population. *Archives of general psychiatry*, 33(9), 1062-1066.

Hjorthøj, C. R., Madsen, T., Agerbo, E., & Nordentoft, M. (2014). Risk of suicide according to level of psychiatric treatment: a nationwide nested case–control study. *Social psychiatry and psychiatric epidemiology*, 49(9), 1357-1365.

Hodgson, R. J., & Rachman, S. (1977). Obsessional-compulsive complaints. *Behaviour research and therapy*, 15(5), 389-395.

Holma, K. M., Melartin, T. K., Holma, I. A., & Isometsä, E. T. (2008). Predictors for switch from unipolar major depressive disorder to bipolar disorder type I or II: a 5-year prospective study. *The Journal of clinical psychiatry*, 69(8), 1267-1275.

House, A. (1989). Hypochondriasis and related disorders: assessment and management of patients referred for a psychiatric opinion. *General Hospital Psychiatry*, 11, 156–165. doi: 10.1016/0163-8343(89)90035-2.

Hurley, A. D. (2006). Mood disorders in intellectual disability. *Current Opinion in Psychiatry*, 19(5), 465-469.

Hyman, S. (2014). Mental health: depression needs large human genetics studies. *Nature.* 515, 189–191. doi: 10.1038/515189°

Iwanami, T., Maeshima, H., Baba, H., Satomura, E., Namekawa, Y., Shimano, T., & Arai, H. (2015). Psychomotor agitation in major depressive disorder is a predictive factor of mood-switching. *Journal of affective disorders*, 170, 185-189.

Jöreskog, K. G., & Sörbom, D. (1986). LISREL VI: Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables, and Least Squares Methods. Mooresville, IN: Scientific Software.

Jöreskog, K. G., & Sörbom, D. (1989). LISREL 7: A Guide to the Program and Applications. Chicago, IL: SPSS.

Jöreskog, K. G., & Sörbom, D. (1993). LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language. Hillsdale, NJ: Scientific Software International.

Jormann, J., Quinn, M. E. (2014) Cognitive processes and emotion regulation in depression. *Depression and Anxiety.* 31: 308–315. pmid:24668779.

Judd, L. L., Schettler, P. J., Akiskal, H., Coryell, W., Fawcett, J., Fiedorowicz, J. G., & Keller, M. B. (2012). Prevalence and clinical significance of subsyndromal manic symptoms, including irritability and psychomotor agitation, during bipolar major depressive episodes. *Journal of affective disorders*, 138(3), 440-448.

Judd, L. L., & Akiskal, H. S. (2003). The prevalence and disability of bipolar spectrum disorders in the US population: re-analysis of the ECA database taking into account subthreshold cases. *Journal of affective disorders*, 73(1), 123-131.

Judd, L. L., Akiskal, H. S., Schettler, P. J., Coryell, W., Maser, J., Rice, J. A., & Keller, M. B. (2003). The comparative clinical phenotype and long term longitudinal episode

course of bipolar I and II: a clinical spectrum or distinct disorders? *Journal of affective disorders*, 73(1), 19-32.

Judd, L. L., Akiskal, H. S., Schettler, P. J., Endicott, J., Maser, J., Solomon, D. A., Leon, A. C., Rice, J. A., & Keller, M. B. (2002). The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Archives of general psychiatry*, 59(6), 530-537.

Kendler, K.S., Eaves, L.J., Walters, E.E., Neale, M.C., Heath, A.C., Kessler, R.C. (1996). The identification and validation of distinct depressive syndromes in a population-based sample of female twins. *Archives of general psychiatry* 53, 391–399

Kilbane, E. J., Gokbayrak, N. S., Galynker, I., Cohen, L., & Tross, S. (2009). A review of panic and suicide in bipolar disorder: does comorbidity increase risk? *Journal of affective disorders*, 115(1), 1-10.

Korszun, A., Moskvina, V., Brewster, S., Craddock, N., Ferrero, F., Gill, M., Jones, I.R., Jones, L.A., Maier, W., Mors, O., Owen, M.J., Preisig, M., Reich, T., Rietschel, M., Farmer, A., McGuffin, P. (2004) . Familiality of symptom dimensions in depression. *Archives of General Psychiatry*, 61, 468 – 474.

Koukopoulos, A., & Koukopoulos, A. (1999). Agitated depression as a mixed state and the problem of melancholia. *Psychiatric Clinics of North America*, 22(3), 547-564.

Koukopoulos, A., & Sani, G. (2014). DSM-5 criteria for depression with mixed features: A farewell to mixed depression. *Acta Psychiatrica Scandinavica*, 129(1), 4-16.

Koukopoulos, A., Reginaldi, D., Tondo, L., Visioli, C., & Baldessarini, R. J. (2013). Course sequences in bipolar disorder: depressions preceding or following manias or hypomanias. *Journal of affective disorders*, 151(1), 105-110.

Koukopoulos, A., Sani, G., Koukopoulos, A. E., Manfredi, G., Pacchiarotti, I., & Girardi, P. (2007). Melancholia agitata and mixed depression. *Acta Psychiatrica Scandinavica*, 115(s433), 50-57

Kraepelin E. (1921) *Manic-depressive insanity and paranoia*. In: Robertson GM, Barclay RBT, editors. XXX. Edinburgh: E & S Livingstone.

Kraepelin E. (1913) *Psychiatrie*. 8th ed. Leipzig: JA Barth.

Krug, S. E., & Laughlin, J. E. (1976). *Handbook for the IPAT depression scale*. Institute for Personality and Ability Testing.

Kupfer, D. J., Frank, E., & Phillips, M. L. (2012). Major depressive disorder: new clinical, neurobiological, and treatment perspectives. *Lancet*, 379 (9820): 1045-1055.

Lancet, T. (2012). Depression and the global economic crisis: is there hope?.

Lazzari, R., & Pancheri, P. (1980). *Questionario di Valutazione dell'Ansia di Stato e di Tratto (State-Trait Anxiety Inventory)*. Firenze: Organizzazioni Speciali.

Lee, S., Guo, W. G., Tsang, A., Mak, A. D. P., Wu, J., & King Lam Ng. (2010). Kathleen Kwok Evidence for the 2008 economic crisis exacerbating depression in Hong Kong. *Journal of Affective Disorders*, 126 (2010) 125–133.

Leverich, G. S., Altshuler, L. L., Frye, M. A., Suppes, T., Keck Jr, P. E., McElroy, S. L., Denicoff, K. D., Obrocea, G., Nolen, W.A., Kupka, R., Walden, J., Grunze, H., Perez, S., Luckenbaugh, D. A., & Post, R. N. (2003). Factors associated with suicide attempts in 648 patients with bipolar disorder in the Stanley Foundation Bipolar Network. *The Journal of clinical psychiatry*, 64(5), 506-515.

Lozano, R., Naghavi, M., & Foreman, K. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, 380: 2095–128.

MacQueen, G., Santaguida, P., Keshavarz, H., Jaworska, N., Levine, M., Beyene, J., & Raina, P. (2017). Systematic review of clinical practice guidelines for failed antidepressant treatment response in major depressive disorder, dysthymia, and subthreshold depression in adults. *The Canadian Journal of Psychiatry*, 0706743716664885.

Magariños, M., Zafar, U., Nissenson, K., and Blanco, C. (2002). Epidemiology and treatment of hypochondriasis. *CNS Drugs*, 16, 9–22. doi: 10.2165/00023210-200216010-00002.

Maj, M., Pirozzi, R., Magliano, L., & Bartoli, L. (2003). Agitated depression in bipolar I disorder: prevalence, phenomenology, and outcome. *American Journal of Psychiatry*, 160(12), 2134-2140.

Manea, L., Gilbody, S., & McMillan, D. (2012). Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. CMAJ 184, E191–E196. doi: 10.1503/cmaj.110829.

Marinagi, C. C., Kaburlasos, V. G., & Tsoukalas, V. T. (2007, October). An architecture for an adaptive assessment tool. In Frontiers In Education Conference-Global Engineering: Knowledge Without Borders, Opportunities Without Passports, 2007. FIE'07. 37th Annual (pp. T3D-11). IEEE.

Maris, R. (1985). The adolescent suicide problem. *Suicide and Life-Threatening Behavior*, 15(2), 91-109.

Marneros, A. (2004). Affective disorders: basic principles regarding clinical course, long-term therapeutic and prophylactic strategies. *Pharmacopsychiatry*, 37(S 2), 148-151.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural equation modeling*, 11(3), 320-341.

McCullough, J. P., & Clark, S. W. (2017). Persistent Depressive Disorder (Dysthymia) and Its Treatment. *Treatments for Psychological Problems and Syndromes*, 153-167.

McGuffin, P., Farmer, A., & Harvey, I. (1991). A polydiagnostic application of operational criteria in studies of psychotic illness: development and reliability of the OPCRIT system. *Archives of general psychiatry*, 48(8), 764-770.

McIntyre, R. S., Filteau, M. J., Martin, L., Patry, S., Carvalho, A., Cha, D. S., Barakat, M., & Miguelez, M. (2014). Treatment-resistant depression: definitions, review of the evidence, and algorithmic approach. *Journal of affective disorders*, 156, 1-7.

McIntyre, R. S., Soczynska, J. K., Cha, D. S., Woldeyohannes, H. O., Dale, R. S., Alsuwaidan, M. T., Gallaugher, L. A., Mansur, R. B., Muzina, D. J., Carvalho, A., & Kennedy, S. H. (2015). The prevalence and illness characteristics of DSM-5-defined "mixed feature specifier" in adults with major depressive disorder and bipolar disorder: Results from the International Mood Disorders Collaborative Project. *Journal of affective disorders*, 172, 259-264.

Merikangas, K. R., Jin, R., He, J. P., Kessler, R. C., Lee, S., Sampson, N. A., & Ladea, M. (2011). Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Archives of general psychiatry*, 68(3), 241-251.

Miller, W. R., & Seligman, M. E. (1975). Depression and learned helplessness in man. *Journal of Abnormal Psychology*, 84 (3): 228-238.

Mitchell, P. B. (2001). The clinical features of bipolar depression: a comparison with matched major depressive disorder patients. *The Journal of clinical psychiatry*, 62(3), 212-216.

Mulin, E., Leone, E., Dujardin, K., Delliaux, M., Leentjens, A., Nobili, F., Dessi, B., Tible, O., Aguera-Ortiz, L., Osorio, R. S., Yessavage, J., Dachevsky, D., Verhey, F., Cruz, A. J., Blanc, O., Llorca, P. M., & Yessavage, J. (2011). Diagnostic criteria for apathy in clinical practice. *International journal of geriatric psychiatry*, 26(2), 158-165.

Murray, C. J., Lopez, A. D., & World Health Organization. (1996). The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020: summary.

National Collaborating Centre for Mental Health (UK. (2010). Depression: the treatment and management of depression in adults (updated edition). British Psychological Society.

Nicassio, P. M., Meyerowitz, B. E., & Kerns, R. D. (2004). The future of health psychology interventions. *Health Psychology*, 23(2), 132.

Nordgaard, J., Sass, L. A., & Parnas, J. (2013). The psychiatric interview: validity, structure, and subjectivity. *European Archives of Psychiatry and Clinical Neuroscience*, 263(4), 353-364.

Nokelainen, P., Silander, T., Tirri, H., Nevgi, A., & Tirri, K. (2001, June). Modeling students' views on the advantages of web-based learning with Bayesian networks. In Proceedings of the Tenth International PEG Conference: Intelligent Computer and Communications Technology–Learning in On-Line Communities (pp. 101-108).

O'Hara, M. W., Neunaber, D. J., and Zekoski, E. M. (1984). Prospective study of postpartum depression: prevalence, course, and predictive factors. *Journal of Abnormal Psychology*, 93, 158–171. doi: 10.1037/0021-843X.93.2.158

Olgiati, P., Serretti, A., & Colombo, C. (2006). Retrospective analysis of psychomotor agitation, hypomanic symptoms, and suicidal ideation in unipolar depression. *Depression and Anxiety*, 23(7), 389-397.

Olin, B., Jayewardene, A. K., Bunker, M., & Moreno, F. (2012). Mortality and suicide risk in treatment-resistant depression: an observational study of the long-term impact of intervention. *PloS One*, 7(10), e48002.

Oquendo, M. A., Waternaux, C., Brodsky, B., Parsons, B., Haas, G. L., Malone, K. M., & Mann, J. J. (2000). Suicidal behavior in bipolar mood disorder: clinical characteristics of attempters and nonattempters. *Journal of affective disorders*, 59(2), 107-117.

Oral, E. T., & Vahip, S. (2004). Bipolar depression: an overview. *IDrugs*, 7(9), 846.

Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical assessment, research & evaluation*, 9(11), 8.

Otto, Y., Andreas, A., von Klitzing, K., Fuchs, S., & Klein, A. M. (2014). Sadness, worries and fears: depression and anxiety disorders in preschool age--results of relevance, symptoms and impairment. *Praxis Der Kinderpsychologie Und Kinderpsychiatrie*, 63(3), 154-176.

Pakriev, S., Vasar, V., Aluoja, A., Saarma, M., & Shlik, J. (1998). Prevalence of mood disorders in the rural population of Udmurtia. *Acta Psychiatrica Scandinavica*, 97(3), 169-174.

Pancheri P, & Carilli L. Standardizzazione di una rating-scale per la valutazione della sintomatologia depressiva. *Rivista di Psichiatria*, 1982; 17: 121–140.

Parker, G., Roy, K., Hadzi-Pavlovic, D., Mitchell, P., Wilhelm, K., Menkes, D. B., Snowdon, J., Loo, C., & Schwitzer, I. (2000). Subtyping depression by clinical features: the Australasian database. *Acta Psychiatrica Scandinavica*, 101(1), 21-28.

Pedrelli, P., Nyer, M., Holt, D., Bakow, B. R., Fava, M., Baer, L., Cassiello, C., Mulligan, M., Cusin, C., & Farabaugh, A. (2013). Correlates of irritability in college students with depressive symptoms. *The Journal of nervous and mental disease*, 201(11), 953.

Perugi, G., Akiskal, H. S., Lattanzi, L., Cecconi, D., Mastrocinque, C., Patronelli, A., Vignoli, S., & Bemi, E. (1998). The high prevalence of "soft" bipolar (II) features in atypical depression. *Comprehensive psychiatry*, 39(2), 63-71.

Perugi, G., Akiskal, H. S., Micheli, C., Toni, C., & Madaro, D. (2001). Clinical characterization of depressive mixed state in bipolar-I patients: Pisa-San Diego collaboration. *Journal of affective disorders*, 67(1), 105-114.

Perugi, G., Micheli, C., Akiskal, H. S., Madaro, D., Socci, C., Quilici, C., & Musetti, L. (2000). Polarity of the first episode, clinical characteristics, and course of manic depressive illness: a systematic retrospective investigation of 320 bipolar I patients. *Comprehensive psychiatry*, 41(1), 13-18.

Perugi, G., Toni, C., Travierso, M. C., & Akiskal, H. S. (2003). The role of cyclothymia in atypical depression: toward a data-based reconceptualization of the borderline–bipolar II connection. *Journal of affective disorders*, 73(1), 87-98.

Petersen, M. A., Groenvold, M., Aaronson, N., Fayers, P., Sprangers, M., & Bjorner, J. B. (2006). Multidimensional computerized adaptive testing of the EORTC QLQ-C30: basic developments and evaluations. *Quality of Life Research*, 15(3), 315-329.

Pettersson, A., Boström, K. B., Gustavsson, P., & Ekselius, L. (2015). Which instruments to support diagnosis of depression have sufficient accuracy? A systematic review. *Nord Journal of Psychiatry*, 69, 497–508. doi: 10.3109/08039488.2015.1008568

Phillips, M. R., Zhang, J., Shi, Q., Song, Z., Ding, Z., Pang, S., Li, X., Zhang, Y., & Wang, Z. (2009). Prevalence, treatment, and associated disability of mental disorders in four provinces in China during 2001–05: an epidemiological survey. *The Lancet*, 373(9680), 2041-2053.

Pitman, A., Krysinska, K., Osborn, D., & King, M. (2012). Suicide in young men. *Lancet*, 379 (9834): 2383-2392.

Plutchik, R., & Van Praag, H. M. (1987). Interconvertability of five self-report measures of depression. *Psychiatry Research*, 22(3), 243-256.

R Core Team (2013). R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. Available online at: http://www.R-project.org/

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied psychological measurement*, 1(3), 385-401.

Rainone, A., and Mancini, F. (2007). *Gli Approcci Cognitivi Alla Depressione*, Vol. 35. Milano: FrancoAngeli.

Raskin, A., Schulterbrandt, J., Reatig, N., & McKEON, J. J. (1969). Replication of factors of psychopathology in interview, ward behavior and self-report ratings of hospitalized depressives. *The Journal of nervous and mental disease*, 148(1), 87-98.

Reed, G. M., Roberts, M. C., Keeley, J., Hooppell, C., Matsumoto, C., Sharan, P., Robles, R., Carvalho, H., Wu, C., Gureje, O., & Leal-Leturia, I. (2013). Mental Health Professionals' Natural Taxonomies of Mental Disorders: Implications for the Clinical Utility of the ICD-11 and the DSM-5. *Journal of clinical psychology*, 69(12), 1191-1212.

Reeves, A., Stuckler, D., McKee, M., Gunnell, D., Chang, S.S., & Basu, S. (2012). Increase in state suicide rates in the USA during economic recession. *Lancet*, 380 (9856): 1813-1814.

Regier, D. A., Kuhl, E. A., & Kupfer, D. J. (2013). The DSM-5: Classification and criteria changes. *World Psychiatry*, 12(2), 92-98.

Reise, S. P., & Waller, N. G. (1990). Fitting the two-parameter model to personality data. *Applied Psychological Measurement*, 14(1), 45-58.

Rihmer, A., Gonda, X., Balazs, J., & Faludi, G. (2008). The importance of depressive mixed states in suicidal behaviour. *Neuropsychopharmacol Hung*, 10(1), 45-49.

Rihmer, Z., & Kiss, K. (2002). Bipolar disorders and suicidal behaviour. *Bipolar Disorders*, 4(s1), 21-25.

Rihmer, Z., & Pestality, P. (1999). Bipolar II disorder and suicidal behavior. Psychiatric *Clinics of North America*, 22(3), 667-673.

Rihmer, Z., Kapitany, B., Gonda, X., & Dome, P. (2013). Suicide, recession, and unemployment. Lancet, 381 (9868): 722-723.

Roberts, K. E., Hart, T. A., & Eastwood, J. D. (2016). Factor structure and validity of the State-Trait Inventory for Cognitive and Somatic Anxiety. *Psychological assessment*, 28(2), 134.

Roca, R. P., Wigley, F. M., & White, B. (1996). Depressive symptoms associated with scleroderma. *Arthritis & Rheumatology*, 39(6), 1035-1040.

Rosellini, R. A., & Seligman, M. E. (1975). Frustation and learned helplessness. Journal of Experimental Psychology: *Animal Behavior Processes*, 1 (2): 149-157.

Rush, A. J., & Beck, A. T. (1978). Cognitive therapy of depression and suicide. *American Journal of Psychotherapy*, 32 (2): 201-219.

Rusch, A., & Wille, R. (1996). Knowledge spaces and formal concept analysis. Data analysis and information systems, 427-436.

Sachs, G. S., Nierenberg, A. A., Calabrese, J. R., Marangell, L. B., Wisniewski, S. R., Gyulai, L., Friedman, E. S., Bowden, C. L., Fossey, M. D., Ostacher, M. J., Ketter, T.A., Patel, J., Hauser, P., Rapport, D., Martinez, J. M., Allen, M. H., Miklowitz, D. J., Otto, M. V., Dennehy, E. B., Thase, M. E. (2007). Effectiveness of adjunctive antidepressant treatment for bipolar depression. *New England Journal of Medicine*, 356(17), 1711-1722.

Sakamoto, S., Kijima, N., Tomoda, A., & Kambara, M. (1998). Factor structures of the Zung Self-Rating Depression Scale (SDS) for undergraduates. *Journal of clinical psychology*, 54(4), 477-487.

Sanavio, E. (2007). *Il primo colloquio nell'assessment clinico. Psicologo: verso la professione*, 1-18.

Sanavio, E., & Sica, C. (2004). *I test di personalità: Inventari e questionari*. Il mulino.

Sanavio, E., & Vidotto, G. (1985). The components of the Maudsley obsessional-compulsive questionnaire. *Behaviour Research and Therapy*, 23(6), 659-662.

Sanavio, E., Bertolotti, G., Michielin, P., Vidotto, G., & Zotti, A. M. (2008). *CBA-2.0 Scale Primarie: Manuale. Una batteria ad ampio spettro per l'assessment psicologico*. Firenze: Organizzazioni Speciali.

Sanavio, E., Bertolotti, G., Michielin, P., Vidotto, G., & Zotti, A. M. (1986). *CBA-2.0 Scale Primarie: Manuale. Una batteria ad ampio spettro per l'assessment psicologico.* Firenze: Organizzazioni Speciali.

Sanavio, E., Bertolotti, G., Michielin, P., Vidotto, G., & Zotti, A. M. (1997). *CBA-2.0 Scale Primarie: Manuale. Una batteria ad ampio spettro per l'assessment psicologico.* (Seconda edizione ampliata ed.). Firenze: Organizzazioni Speciali.

Sani, G., Vöhringer, P. A., Napoletano, F., Holtzman, N. S., Dalley, S., Girardi, P., Ghaemi, S. N., & Koukopoulos, A. (2014). Koukopoulos′ diagnostic criteria for mixed depression: A validation study. *Journal of affective disorders*, 164, 14-18.

Santor, D. A., Gregus, M., & Welch, A. (2006). FOCUS ARTICLE: Eight decades of measurement in depression. *Measurement: Interdisciplinary Research and Perspectives*, *4*(3), 135-155.

Sato, T., Bottlender, R., Schröter, A., & Möller, H. J. (2003). Frequency of manic symptoms during a depressive episode and unipolar 'depressive mixed state'as bipolar spectrum. *Acta Psychiatrica Scandinavica*, 107(4), 268-274.

Schildkraut, J. J., & Kety, S. S. (1967). Biogenic amines and emotion. Science, 156 (3771):21-37.

Schloesser, R. J., Orvoen, S., Jimenez, D. V., Hardy, N. F., Maynard, K. R., Sukumar, M., Manji, H. K., Gardier, A. M., David, D. J., & Martinowich, K. (2015). Antidepressant-like effects of electroconvulsive seizures require adult neurogenesis in a neuroendocrine model of depression. *Brain stimulation*, 8(5), 862-867.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53–61. doi: 10.1007/BF02289209.

Seligman, M. E. (1972). Learned helplessness. Annual Review of Medicinal, 23, 407-412.

Seligman, M. E. (1978). Learned helplessness as a model of depression. Comment and integration. *Journal of Abnormal Psychology*, 87 (1), 165-179.

Seligman, M. E., Weiss, J., Weinraub, M., & Schulman, A. (1980). Coping behavior: learned helplessness, physiological change and learned inactivity. *Behaviour Research and Therapy*, 18 (5): 459-512.

Serra, F., Spoto, A., Ghisi, M., & Vidotto, G. (2015a). Formal Psychological Assessment in evaluating depression: a new methodology to build exhaustive and irredundant adaptive questionnaires. *PloS One*, 10(4), e0122131.

Serra, F., Spoto, A., Ghisi, M., & Vidotto, G. (2017). Improving Major Depressive Episode Assessment: A New Tool Developed by Formal Psychological Assessment. *Frontiers in psychology*, 8.

Serra, F., Spoto, A., Vidotto, G. (2015b). A formal approach to assist clinical evaluation of psychological disorders. A procedure based on "items content" – "diagnostic criteria" relations.4th ANNUAL INTERNATIONAL CONFERENCE - PROCEDINGS (CBP 2015), 118-123.

Serra, G., & Fratta, W. (2007). A possible role for the endocannabinoid system in the neurobiology of depression. *Clinical Practice and Epidemiology in Mental Health*, 19, 3-25.

Serra, G., Demontis, F., Serra, F., De Chiara, L., Spoto, A., Girardi, P., Vidotto, G., & Serra, G. (2014). Memantine: New prospective in bipolar disorder treatment. *World journal of psychiatry*, 4(4), 80.

Serretti, A., & Olgiati, P. (2005). Profiles of "manic" symptoms in bipolar I, bipolar II and major depressive disorders. *Journal of affective disorders*, 84(2), 159-166.

Shapiro, M. B. (1951). An experimental approach to diagnostic psychological testing. *Journal of Mental Science*, 97, 748–764. doi: 10.1192/bjp.97.409.748

Sharma, R., & Markar, H. R. (1994). Mortality in affective disorder. *Journal of affective disorders*, 31(2), 91-96.

Sheehan, D., V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., & Dumbar, G. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, 59, 22–33.

Simms, L., Goldberg, L., Roberts, J., Watson, D., Welte, J., & Rotterman, J. (2011). Computerized adaptive assessment of personality disorder: Introducing the CAT-PD project. *Journal of Personality Assessment*, 93, 380-389.

Singh, S. M., & Sharma, A. (2013). The prevalence and correlates of guilt in depression: a study from North India. *Asian journal of psychiatry*, 6(6), 622-623.

Smits, N., Finkelman, M. D., & Kelderman, H. (2016). Stochastic curtailment of questionnaires for three-level classification shortening the CES-D for assessing low, moderate, and high risk of depression. *Applied Psychological Measurement*, 40, 22-36.

Sobin, C., & Sackeim, H. A. (1997). Psychomotor symptoms of depression. *The American journal of psychiatry*, 154(1), 4.

Spiegel, R., & Nenh, Y. P. (2004). An expert system supporting diagnosis in clinical psychology. WIT Transactions on Information and Communication Technologies, 31.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *STAI Manual for the Stait-Trait Anxiety Inventory ("self-evaluation Questionnaire")*. Palo Alto, Calif.: Consulting Psychologists Press.

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research diagnostic criteria: rationale and reliability. *Archives of general psychiatry*, 35(6), 773-782.

Spoto, A., Bottesi, G., Sanavio, E., & Vidotto, G. (2013). Theoretical foundations and clinical implications of formal psychological assessment. *Psychotherapy and psychosomatics*, 82(3), 197-199.

Spoto, A. (2011). Formal Psychological Assessment Theoretical and Mathematical Foundations. PhD thesis.

Spoto, A., Stefanutti, L., & Vidotto, G. (2010). Knowledge space theory, formal concept analysis, and computerized psychological assessment. *Behavior research methods*, 42(1), 342-350.

SPSS, I. (2012). Statistical package for the social sciences. International Business Machines Corporation SPSS Statistics, Armonk, NY, USA.

Stefánsson, J. G., Lindal, E., Björnsson, J. K., & Guðomundsdóttir, Á. (1991). Lifetime prevalence of specific mental disorders among people born in Iceland in 1931. *Acta Psychiatrica Scandinavica*, 84(2), 142-149.

Subica, A. M., Fowler, J. C., Elhai, J. D., Frueh, B. C., Sharp, C., Kelly, E. L., & Allen, J. G. (2014). Factor structure and diagnostic validity of the Beck Depression Inventory–II with adult clinical inpatients: Comparison to a gold-standard diagnostic interview. *Psychological assessment*, 26(4), 1106.

Sullivan, P. F., Kessler, R. C., & Kendler, K. S. (1998). Latent class analysis of lifetime depressive symptoms in the national comorbidity survey. *American Journal of Psychiatry*, 155(10), 1398-1406.

Swann, A. C. (2013). Activated depression: mixed bipolar disorder or agitated unipolar depression? *Current psychiatry reports*, 15(8), 1-8.

Swann, A. C., Secunda, S. K., Katz, M. M., Croughan, J., Bowden, C. L., Koslow, S. H., Berman, N., & Stokes, P. E. (1993). Specificity of mixed affective states: clinical comparison of dysphoric mania and agitated depression. *Journal of affective disorders*, 28(2), 81-89.

Tafet, G. E., Idoyaga-Vargas, V. P., Abulafia, D. P., Calandria, J. M., Roffman, S. S., Chiovetta, A., & Shinitzky, M. (2001). Correlation between cortisol level and serotonin uptake in patients with chronic stress and depression. *Cognitive, Affective, & Behavioral Neuroscience*, 1(4), 388-393.

Takeshima, M., & Oka, T. (2013). Association between the so-called "activation syndrome" and bipolar II disorder, a related disorder, and bipolar suggestive features in outpatients with depression. *Journal of affective disorders*, 151(1), 196-202.

Treadway, M. T., & Zald, D. H. (2011). Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neuroscience & Biobehavioral Reviews*, 35(3), 537-555.

Tsuboi, K., & Masuko, M. (2001). Panic disorder and antidepressants. Nihon rinsho. *Japanese journal of clinical medicine*, 59(8), 1592-1598.

Tsuno, N., Besset, A., & Ritchie, K. (2005). Sleep and depression. *The Journal of clinical psychiatry*, 66(10), 1254-1269.

Uher, R., Payne, J. L., Pavlova, B., & Perlis, R. H. (2014). Major depressive disorder in dsm-5: implications for clinical practice and research of changes from DSM-IV. *Depression and Anxiety*, 31(6), 459-471

Van Zaane, J., Van den Berg, B., Draisma, S., Nolen, W. A., & Van den Brink, W. (2012). Screening for bipolar disorders in patients with alcohol or substance use disorders: performance of the mood disorder questionnaire. *Drug and Alcohol Dependence*, 124 (3): 235-241.

Vázquez, G. H., Tondo, L., Undurraga, J., & Baldessarini, R. J. (2013). Overview of antidepressant treatment of bipolar depression. *International Journal of Neuropsychopharmacology*, 16(7), 1673-1685.

Vieta, E., Gasto, C., Otero, A., Nieto, E., & Vallejo, J. (1997). Differential features between bipolar I and bipolar II disorder. *Comprehensive psychiatry*, 38(2), 98-101.

Vöhringer, P. A., & Perlis, R. H. (2016). Discriminating Between Bipolar Disorder and Major Depressive Disorder. *Psychiatric Clinics of North America*, 39(1), 1-10.

Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., & Coggeshall, M. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, 388(10053), 1545.

Wainer, H. (2000). CATs: Whither and whence. *ETS Research Report Series*, 2000(2).

Waraich, P., Goldner, E. M., Somers, J. M., & Hsu, L. (2004). Prevalence and incidence studies of mood disorders: a systematic review of the literature. *The Canadian Journal of Psychiatry*, 49(2), 124-138.

Weiss, D. J. (2004). Computerized adaptive testing for effective and efficient measurement in counseling and education. *Measurement and Evaluation in Counseling and Development*, 37(2), 70.

Weissman, M. M., Bland, R. C., Canino, G. J., Faravelli, C., Greenwald, S., Hwu, H. G., Joyce, P. R., Karam, E. G., Lee, C. K., Lellouch, J., Lépine, J. P., Newman, S.C., Rubio-Stipec, M., Wells, J. E., Wickramaratne, P. J., Wittchen, H. U., Yeh, U. K., & Lépine, J. P. (1996). Cross-national epidemiology of major depression and bipolar disorder. *Jama*, 276(4), 293-299.

Wille, R. (1982). "Restructuring lattice theory: an approach based on hierarchies of concepts," in Ordered Sets, ed I. Rival (Dordrecht: Reidel), 445–470.

Wing, J. K., Babor, T., Brugha, T., Burke, J., Cooper, J. E., Giel, R., Jablenschi, A., Regier, D., & Sartorius, N. (1990). SCAN: Schedules fonr Clinical Assessment in Neuropsychiatry. *Archives of general psychiatry*, 47(6), 589-593.

Wittchen, H. (2000). Epidemiology of affective disorders. Helmchen H, Henn F, Lauter H, Sartorius N: Contemporary Psychiatry.

Wolpe, J., & Lang, P. J. (1964). A fear survey schedule for use in behaviour therapy. *Behaviour Research and Therapy*, 2(1), 27-30.

World Health Organization. (2017). Depression and other common mental disorders: global health estimates.

Wright, J. G., & Feinsten, A. R. (1992). A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *Journal of clinical epidemiology*, 45(11), 1201-1218.

Yong, S., Rambli, A., Rohaya, D., & Anh, N. (2007, October). Depression consultant expert system. Paper presented at the 6th annual seminar on science and technology, Tawau, Sabah, Malaysia.

Zimmerman, M., McGlinchey, J. B., Chelminski, I., & Young, D. (2012). Diagnostic co-morbidity in 2300 psychiatric out-patients presenting for treatment evaluated with a semi-structured diagnostic interview. *Psychological and Medicine*, 38 (2): 199-210.

Zung G. A self-rating depression scale. *Archives of General Psychiatry*. (1965). 12: 63–70. pmid:14221692.