The glue of (ab)normal mental life: Networks of interacting thoughts, feelings and behaviors

Cramer, A.O.J.

Publication date
2013
Document Version
Final published version

The glue of abnormal mental life

Networks of interacting thoughts, feelings and behaviors

The glue of abnormal mental life: Networks of interacting thoughts, feelings and behaviors

Angélique O.J. Cramer

The glue of (ab)normal mental life: Networks of interacting thoughts, feelings and behaviors

Angélique Cramer

Cramer, Angélique Odette Joanne
The glue of (ab)normal mental life: Networks of interacting thoughts, feelings and behaviors
Ph.D. Thesis University of Amsterdam. - Met samenvatting in het Nederlands.
ISBN: 978-94-6191-810-9
printed by: Ipskamp Drukkers
Cover designed by Tone Kristen Bjordam. The cover is a still image from the video project "Critical Transitions" made in collaboration with Marten Scheffer.

This thesis is dedicated to my beloved father,
Ronald Cramer (1941-2008),
I miss your sense of humor, advice, and friendship every day

THE GLUE OF (AB)NORMAL MENTAL LIFE: NETWORKS OF INTERACTING THOUGHTS, FEELINGS AND BEHAVIORS

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom,
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 6 september 2013, te 14:00 uur

door

Angélique Odette Joanne Cramer

geboren te Heemstede

Promotiecommissie

Promotor 1	Prof. Dr. D. Borsboom
Promotor 2	Prof. Dr. H.L.J. van der Maas
Overige leden:	Prof. Dr. J.H. Kamphuis
	Prof. Dr. M. Scheffer
	Prof. Dr. F. Tuerlinckx
	Dr. S. van der Sluis
	Prof. Dr. E.-J. Wagenmakers
	Dr. M. Wichers

Faculteit der Maatschappij- en Gedragswetenschappen

Contents

Chapter 1

Introduction

Major depression is caused by a gene, which causes serotonin shortage in the brain. True or false? Likewise: extraversion causes party-going in individuals. Again: true or false? These and similar statements are not a figment of my imagination, these are real statements in either the scientific literature or in the popular press; and these statements depict a deceptively simple picture of what psychological constructs such as major depression and extraversion are. Small and inconsequential as such statements may seem, they speak volumes of how people—scientists and laypeople alike—(implicitly) think about psychological constructs in general: for example, 1) they are ultimately reducible to specific (neuro)biological properties (e.g, disorder X is caused by gene A), 2) they operate in the minds of individual people (e.g., personality trait Y is somewhere in my brain), and 3) we know for a fact how these constructs, in your and in my mind, operate to cause a constellation of (pathological) thoughts, feelings, and behaviors (e.g., we know how disorder X causes symptoms B and C). One of the main forces that drove the inception of this dissertation was the realization that these statements are most likely false. Let us consider some of the facts: despite tremendous efforts to find it, there is no gene (or constellation of genes) that explains more than a fraction of the phenotypic variance in major depression, or any other psychological construct (Kendler, 2005a; Stefanis, 2008); antidepressants, which aim to augment available serotonin levels in the brain, do not work in all patients with major depression (e.g., Lacasse & Leo, 2005); there is no single experiment that has ever shown how exactly extraversion causes party-going behavior in individual people (Cramer, van der Sluis, et al., 2012); etc. That is, given the actual findings, it is at best premature to claim to know what psychological constructs are and how they are caused, let alone presupposing that biological reductionism should be the ultimate goal of psychological science. At worst, we are essentially clueless regarding the nature of psychological constructs.

There is one thing that we do know, which forms the starting point of this dissertation: undisputed and consistent, perhaps the number one fact in clinical

Partly adapted from: Cramer, A. O. J., & Borsboom, D. (under review). Problems attract problems: The network perspective on mental disorders.

Borsboom, D., & Cramer, A. O. J. (in press). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology.*

and personality psychology is that some (pathological) thoughts, feelings, and behaviors co-occur more often with one another than with other thoughts, feelings and behaviors. For example, liking parties and liking people tend to co-occur in individuals while liking parties and sorting one's socks by color do not. Likewise in psychiatry, since the early 20th century, psychiatrists such as Kraepelin (1923) and Lewis (1934) have observed that, for example, depressed mood and feelings of guilt tended to co-occur more frequently with one another than, say, depressed mood and having panic attacks. Now, the number one question is, or should be, what causes these specific patterns of covariation? What is the glue that holds together some (pathological) thoughts, feelings and behaviors?

For decades the answer has been: different underlying causes. That is, in personality psychology, the dominant idea is that, for instance, liking parties and liking people co-occur because they share the same underlying cause, namely extraversion. Liking parties and sorting one's socks by color, in contrast, covary to a lesser extent because they have different underlying causes: liking parties is caused by extraversion while sorting socks by color is caused by conscientiousness. Likewise in psychiatry, depressed mood and feelings of guilt co-occur frequently because they are caused by the same underlying disorder: major depression. In contrast, depressed mood and panic attacks do not co-occur as frequently because they are caused by different disorders: depressed mood is caused by major depression while panic attacks are caused by panic disorder. In particular for mental disorders, the idea of disorders being *common causes* of their symptoms probably has its roots in the successful paradigm of Western medicine (Hyland, 2011). For quite some medical diseases, it makes sense to postulate that disorders cause their respective symptoms: for example, a lung tumor that, because of its physical presence in someone's lungs, causes shortness of breath, chest pains and coughing up blood. This medical disease model is exactly what has fueled the quest for the analogy of the 'tumor' in the case of mental disorders: some pathophysiological correlate that, like a lung tumor, causes the symptoms of a particular mental disorder. For instance, in the case of major depression, a plethora of work appears to show that major depression is associated with a host of pathophysiological correlates: e.g., serotonin depletion, allelic variants of certain genes that appear to predict treatment outcome (Ogilvie et al., 1996; Serretti, Kato, De Ronchi, & Kinoshita, 2007; Wong, Dong, Andreev, Arcos-Burgos, & Licinio, 2012), and atrophy in brain areas such as the hippocampus (MacQueen & Frodl, 2011; Sheline, Wang, Gado, Csernansky, & Vannier, 1996).

The problem, however, with many of these findings is at least fourfold: 1) *specificity*: abnormalities in the serotonin reuptake function, for example, are not only implicated in the etiology of major depression but in that of obsessive-compulsive disorder, substance abuse and anxiety disorders (Nakamura, Ueno, Sano, & Tanabe, 2000); 2) *explained variance*: when combining all possible candidate genetic variants, they still explain only a very small portion of the variance in major depression (Wong et al., 2012); 3) *cause or effect*: for hippocampal atrophy, for example, it is not clear whether this is a cause or effect of (repeated episodes of) major depression (MacQueen et al., 2003); and 4) *no omnipresence*: serotonin depletion, for example, is not present in a substantial proportion of patients with major depression (Lacasse & Leo, 2005). As such, given these problems, the com-

mon causes that correspond to mental disorders (the 'tumors') either do not exist or else are *very* hard to find. There are two ways in which one can respond to such a gap between theory and empirical evidence. One way, the road that has usually been taken in the past years, is that we should look harder. With more participants, better research equipment and ever more intricate ways of analyzing the data, we will eventually find the 'tumor' equivalent and its associated physiological and genetic abnormalities. The other road, and the one we, and others (McGrath, 2005; Kendler, Zachar, & Craver, 2011) have taken, is to accept these findings as an indication that we may need to rethink the nature of mental disorders, and of psychological constructs in general.

So if not common causes, what then is the glue of (ab)normal mental life? In this dissertation, I will argue that direct interactions between (pathological) thoughts, feelings and behaviors in a network are the glue. As such, liking parties and liking people do not covary because they are both caused by extraversion; they covary because they directly interact with one another: e.g., enjoying the company of other people prompts a person to seek out environments in which one can mingle with people, for example at parties. Likewise, depressed mood and feelings of guilt do not co-occur frequently because they are both caused by major depression; they covary because of a direct relation: e.g., depressed mood causes a person to feel guilty towards friends and family for being so blue all the time. In Chapter 2, this *network approach* will be explicated and exploratively tested for mental disorders, in particular for comorbidity between major depression and generalized anxiety disorder. Also, the network model will be conceptually contrasted to the mathematical formalization of the common cause idea, namely latent variable models. In Chapter 3, I respond to various commentaries that were written in reply to Chapter 2. In Chapter 4, I will show that the network model explains the phenomenon that various stressful life events (e.g., the loss of a loved one) influence different symptoms of major depression; a finding that is not easily accommodated by latent variable models. In Chapter 5, I present a mathematical formalization of the network model in which the probability of a symptom becoming activated is a logistic function of the activation of its neighboring symptoms in the network. Additionally, I show that this model is able to explain empirical phenomena such as spontaneous recovery, and that the model accommodates taxonomic as well as continuous views on major depression. In Chapter 6, I outline a network perspective on normal personality traits and show how taking this perspective might change our outlook on the concept of a 'trait', the relationship between genes and traits, and between personality and psychopathology. In Chapter 7, I respond to various commentaries that were written in reply to Chapter 6. Finally, Chapter 8 provides a practical guide to construct and analyze networks.

Chapter 2

Comorbidity: A network perspective

Abstract

The pivotal problem of comorbidity research lies in the psychometric foundation it rests on, that is, *latent variable theory*, in which a mental disorder is viewed as a latent variable that *causes* a constellation of symptoms. From this perspective, comorbidity is a (bi)directional relationship between multiple latent variables. We argue that such a latent variable perspective encounters serious problems in the study of comorbidity, and offer a radically different conceptualization in terms of a *network approach*, where comorbidity is hypothesized to arise from direct relations between symptoms of multiple disorders. We propose a method to visualize comorbidity networks and, based on an empirical network for major depression and generalized anxiety, we argue that this approach generates realistic hypotheses about pathways to comorbidity, overlapping symptoms, and diagnostic boundaries, that are not naturally accommodated by latent variable models: Some pathways to comorbidity through the *symptom space* are more likely than others; those pathways generally have the same direction (i.e., from symptoms of one disorder to symptoms of the other); overlapping symptoms play an important role in comorbidity; and boundaries between diagnostic categories are necessarily fuzzy.

Adapted from: Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences, 33*, 137-150.

If suffering from a single mental disorder is bad, suffering from multiple mental disorders (i.e., *comorbidity*) is worse. Compared to suffering from a single mental disorder, comorbidity is consistently associated with a greater demand for professional help, a poorer prognosis, greater interference with everyday life, and higher suicide rates (e.g., U. Albert, Rosso, Maina, & Bogetto, 2008; T. A. Brown, Antony, & Barlow, 1995; Schoevers, Deeg, van Tilburg, & Beekman, 2005). Also, among people who meet diagnostic criteria for one mental disorder, approximately 45% receive additional diagnoses (e.g., Kessler, Chiu, Demler, & Walters, 2005). Thus, comorbidity is a widespread and serious problem, the underpinnings of which need to be unraveled. Indeed, the comorbidity issue has been studied extensively in the past decades (e.g., Anderson, Williams, McGee, & Silva, 1987; Angold, Costello, & Erkanli, 1999; Boyd et al., 1984; T. A. Brown, Campbell, Lehman, Grisham, & Mancill, 2001; Kashani et al., 1987; Kessler et al., 1994, 2004; Kessler, Berglund, Demler, Jin, & Walters, 2005; Low, Cui, & Merikangas, 2008; Merikangas et al., 1998; Moffitt et al., 2007; Neale & Kendler, 1995).

However, although considerable progress towards furthering our understanding of comorbidity has been made, some pivotal questions remain unanswered. Probably the most crucial question is *what* we observe when two disorders covary: a genuine phenomenon that is independent of our diagnostic criteria, measurement scales, and measurement models, or (in part) an artifact of the structure of these criteria and models (e.g., see Borsboom, 2002; Neale & Kendler, 1995)? The former possibility holds that a genuine source of comorbidity rates exists. As such, the disorders *themselves* are comorbid, which *causes* the symptoms of such comorbid disorders to correlate. The latter possibility holds that comorbidity is produced by the way we empirically identify these disorders; for instance, because disorders often share a number of symptoms, which leads to an artificially increased comorbidity rate. Thus, in this view, comorbidity is an artifact of the diagnostic system.

In this chapter, we argue that these possibilities are not exhaustive. Specifically, we argue that comorbidity is not an artifact. However, we do contend that comorbidity, as it has been studied so far, *is* dependent on the way we psychometrically portray disorders and comorbidity between them: namely, with a latent variable model (e.g., factor models, item response models). Within this psychometric framework, comorbidity is generally conceptualized as a (bi)directional relationship between two latent variables (i.e., disorders) that underlie a set of symptoms. In our view, there are good reasons to doubt the validity of the psychometric assumptions that underlie this approach. We discuss these reasons and propose an alternative conceptualization of the relation between symptoms and disorders that offers a natural way of explaining comorbidity

The central idea is that disorders are *networks* that consist of *symptoms* and *causal* relations between them. In a nutshell, what binds, say, the set of depression symptoms, is that they are thus connected through a dense set of strong causal relations. With regard to comorbidity, such a *network approach* presents a radically different conceptualization of comorbidity, in terms of direct relations between the symptoms of multiple disorders.

In contrast to existing perspectives, it is inappropriate to say that the symptoms *measure* the disorder in question. The reason is that the presence of direct causal relations between symptoms contradicts the essential assumptions that underlie psychology's main class of measurement models (latent variable models; e.g. Borsboom, 2005, 2008b). In fact, a network approach nullifies the need to invoke latent variables as an explanation of the covariance between symptoms. In a network approach, the relation between symptoms and disorders (or, more generally, test scores and constructs) should not be viewed as one of *measurement*, but as one of *mereology*: The symptoms do not measure the disorder, but are part of it (see also Markus, 2008 for a discussion of the role of mereology and causality in statistical modeling). This is consistent with McGrath's (2005)

observation that theoretical terms in psychology, such as 'depression', may often refer to complex constellations of variables, rather than to a single latent structure.

Hence, it is likely that comorbidity's true colors are obscured by methodological problems that spring from the assumptions underlying current techniques. The specifics of those problems vary, but all bear one striking resemblance: they are at least in part attributable to the notion that one can focus on *diagnoses* in current comorbidity research, because diagnoses serve as reliable proxies for the latent variables that supposedly underlie them. In this chapter, we provide an in-depth discussion of these problems and show that the network approach avoids them.

The structure of this chapter is as follows. First, we introduce the network approach by contrasting it to the latent variable variable model. We subsequently propose an integrative way to visualize comorbidity as a symptom network, and discuss the basic features of an empirical network for major depressive disorder (MDD) and generalized anxiety disorder (GAD), based on data from the National Comorbidity Survey Replication[1]. Then, we discuss three additional methodological problems that characterize current comorbidity research and argue that adopting a network approach may help in answering questions that are, in our view, crucial when painting an accurate picture of comorbidity: How important are symptoms that overlap between two disorders as sources of comorbidity? Can we identify symptoms of a disorder that put someone at more risk of developing a second disorder compared to other symptoms? Is there an order in which people generally develop one particular disorder first and another disorder second?

Mental disorders: Networks of directly related symptoms instead of latent variables

Measurement models used in clinical and personality research have one thing in common: the assumption that there is some attribute we cannot observe directly (i.e., is "latent")—MDD or extraversion, for instance—and therefore, must be *measured indirectly* through the presence or absence of certain observable variables (e.g., MDD is measured by depressed mood and extraversion is measured by party-going behavior; McCrae & Costa, 2008; see Michell, 2005 for a detailed explanation of measurement in science). In doing so, *latent variable models* are consistent with the hypothesis that the latent attribute has causal relevance for the observed values of symptoms (e.g., see Borsboom, Mellenbergh, & van Heerden, 2003, 2004; Borsboom, 2008b; Hood, 2008). In this view, for instance, depression (i.e., the latent attribute) *causes* the occurrence of symptoms such as fatigue.

In line with this idea, it is commonly hypothesized that comorbidity arises due to some direct relation between two latent variables; for example, a substantial correlation as depicted in Figure 2.1 (e.g., MDD and GAD; Neale & Kendler, 1995). Some theorize even further, and hypothesize that a direct relation between two latent variables actually reflects the existence of a 'super disorder' —for example, in models in which the super disorder 'negative affect' causes a variety of mental disorders (e.g., depression), which, in

[1]The National Comorbidity Survey Replication (NCS-R) is a nationally representative household survey of English speakers 18 years and older in the United States (see Kessler et al., 2004). The NCS-R survey schedule is the version of the World Health Organization (WHO) Composite International Diagnostic Interview that is developed for the WHO World Mental Health Survey Initiative (WMH-CIDI; Kessler & Ustun, 2004). The interviews were conducted between February 2001 and April 2003. A total of 9282 respondents participated in Part 1 of the interview (core diagnostic assessment) that we used for this chapter. The symptoms that participants reported within one disorder all occurred within the same time frame.

turn, cause observable symptoms (e.g., see Barlow, Allen, & Choate, 2004). In accordance with both views on comorbidity, current comorbidity research mainly focuses on diagnoses as proxies of the latent disorders and computes tetrachoric or odds ratios between those proxies. Although this methodology has yielded important insights (e.g., T. A. Brown et al., 2001; Kessler et al., 1994, 2005; Merikangas et al., 1998; Moffitt et al., 2007), the latent variable model may not always offer the best psychometric perspective to conceptualize mental disorders (see also Borsboom, 2008b).

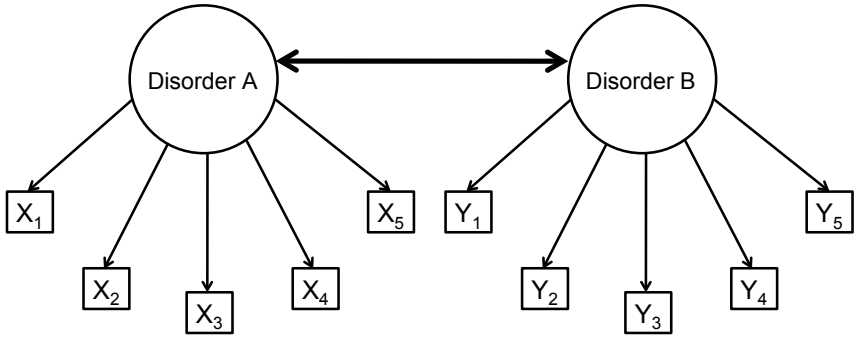

Figure 2.1: A model of comorbidity between disorders A and B under the standard assumptions of latent variable modeling. The *circles* represent the disorders (i.e., latent variables) and the *rectangles* represent the observable core symptoms of those disorders (i.e., $X_1 - X_5$ for disorder A, and $Y_1 - Y_5$ for disorder B). In this model, comorbidity is viewed as a correlation between the latent variables, visualized by the *thick bidirectional edge* between disorders A and B.

To see this, it is useful to consider the essence of latent variable modeling, the *common cause* hypothesis, in more detail. The common cause hypothesis posits that a latent variable causes its observable indicators. If one adopts this hypothesis for a particular set of variables, then one has to accept an important consequence: *The observable indicators cannot be directly related*; that is, if a single common cause is held responsible for the occurrence of a particular set of variables, then covariation between those variables is entirely attributable to the common cause. It is important to note here that we are referring to the *psychometric* as opposed to a *clinical* interpretation of a latent variable model. In the clinical interpretation, clinicians adhere to the existence of a latent variable while at the same time acknowledging direct relations between symptoms. In a strict psychometric sense, a latent variable model does not allow for many direct relations since the majority of covariance between symptoms needs to be explained by the common cause. As such, psychometric latent variable models imply that correlations between observable indicators are, in a non-trivial sense, spurious. When statistically modeling the relationship between a hypothesized latent variable and a set of indicators, the fact that the indicators cannot be directly related results in the statistical assumption of *local independence* (such assumptions are made, for instance, in the models used in Aggen, 2005; C. A. Hartman et al., 2001; Krueger, 1999): when fitting a latent variable model to observed data, any two indicators are conditionally independent given the latent variable (Lord & Novick, 1968). As such, local independence is a statistical consequence of adopting the hypothesis that a common cause structure gave rise to the associations in the data.

In our view, a common cause structure is unlikely to hold for symptoms of mental disorders. For instance, consider 'sleep disturbances' and 'fatigue', both of which are DSM-IV symptoms of MDD (see *Diagnostic and Statistical Manual of Mental Disorders, 4th edition*; APA, 1994). If one adopts the common cause hypothesis, a high positive

correlation between these symptoms is entirely due to the common influence of the latent variable, MDD. It is questionable whether this is plausible. For instance, a direct causal relationship between those symptoms is likely to hold in at least a subset of people who experience them: If you don't sleep, you get tired. Another example: Is it plausible to assume that GAD necessarily causes both chronic worry and a difficulty to concentrate? It may well be that a direct causal relationship exists between these symptoms: the more you worry, the more difficult it is to concentrate at other things.

Thus, it appears likely that latent variable models do not optimally conceptualize the relationship between mental disorders and their symptoms. This is not to say we object to the notion that symptoms of various disorders tend to cluster together in predictable ways and that, as such, disorders may be pragmatically useful to denote such clusters (see C. A. Hartman et al., 2001). However, we do suggest that mental disorders may not explain covariation between symptoms in the way a latent variable pictures the situation. If this is so, then even though the application of latent variable modeling may have considerable instrumental utility (e.g., in facilitating predictions or gauging rough differences between people), one cannot plausible say that the symptoms actually *measure* a latent variable. Therefore, we consider it important to examine relationships between individual symptoms more closely.

Initiating such an endeavor is a major goal of this chapter. As a starting point, we propose to use the theory of complex networks. This theory has provided major contributions to current knowledge about the structure of the World Wide Web, power grids, and neural systems (e.g., see Albert & Barábasi, 1999, 2002; Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006; Strogatz, 2001; X. F. Wang, 2002). The basic idea of the network approach is straightforward: We define and analyze relationships between symptoms, without assuming *a priori* that such relationships arise from a mental disorder as a common cause (Borsboom, 2008b; van der Maas et al., 2006). Simply put, in such a network, a disorder is conceptualized as a cluster of directly related symptoms. In a fairly recent study, Kim and Ahn (2002) showed that this conceptualization comes naturally to some clinicians: depression, anorexia nervosa, antisocial personality disorder, and specific phobia were all characterized as clusters of causally related symptoms. And, adhering to such a network perspective cannot be reconciled with the psychometric properties of a latent variable model. Thus, when modeling comorbidity, we no longer assume a direct relation between two latent variables. Instead, we model comorbidity in terms of a set of distinct relationships between symptoms of distinct disorders.

A network model represents symptoms as *nodes* in a graph and the relationships between them as *edges*. Figure 2.2 depicts an example of such a graph for two disorders: two sets of symptoms belong to two distinct mental disorders. Within each disorder, all symptoms are connected with one another, but between disorders, there are fewer (or weaker) edges between the symptoms. There are also symptoms that do not clearly belong to one or the other disorder, because they receive and send out effects to the symptoms in *both* of the disorders (i.e., overlapping symptoms). If such symptoms overlap perfectly, they can be collapsed into a single symptom, which we propose to call a *bridge symptom*. We hypothesize that in clinical practice, such bridge symptoms turn up as symptoms that are used in diagnostic schemes, such as the DSM-IV, for multiple disorders.

Our hypothesis regarding the crucial role of bridge symptoms in explaining comorbidity can be tested, just as a host of hypotheses can be tested with latent variable models. For binary data, a statistical parameterization of the network is a loglinear model, which is implemented in the gRbase package for R (Dethlefsen & Hojsgaard, 2005). In short, with a loglinear model, one searches for the most parsimonious model—among models ranging from only main effects through model with nth-order interactions—that

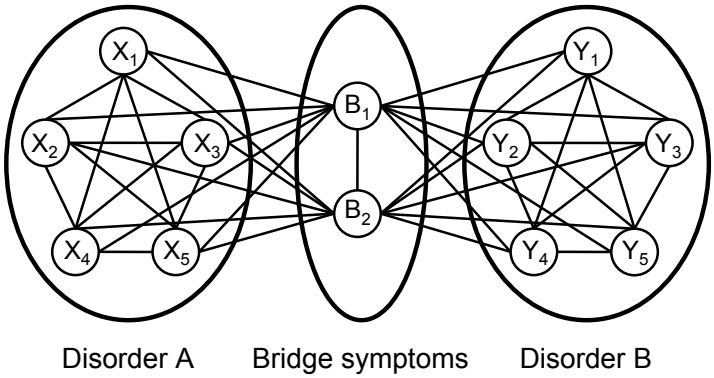

Figure 2.2: Comorbidity under a network approach. Disorder A consists of bidirectionally related symptoms X1-X5, and disorder B consists of symptoms Y1-Y5. Symptoms B1 and B2 are *bridge symptoms* that overlap between disorders A and B. In this model, comorbidity arises as a result of direct relations between the bridge symptoms of two disorders.

accounts for the distribution of cases in contingency tables of categorical variables (e.g., see Agresti, 2002). If the main effects model should turn out to be the best model, then the MDD and GAD symptoms are statistically independent, and our hypothesized bridge model should be rejected accordingly. Thus, in gRbase, we fitted a model like the one shown in Figure 2.2 to the NCS-R MDD and GAD data: All symptoms of MDD/GAD, including the bridge symptoms, are connected with one another, and comorbidity arises only through connections between overlapping symptoms, on the one hand, and other symptoms of MDD/GAD, on the other hand[2]. We used the Akaike Information Criterion (AIC) to compare the fit of three models: (1) with only main effects, (2) with first-order interactions within disorders (including bridge symptoms, as in Figure 2.2), and (3) with second-order interactions within disorders (including bridge symptoms). Of these three models, the best fitting model according to the AIC is the one with first-order interactions (AIC differences are: (2) - (1) = -177.551 and (3) - (2) = 347.123). Thus, according to this analysis, the bridge model holds with all variables being statistically dependent on one another. Naturally, such a single fit is not sufficient to conclude that this model is the best choice, especially since—considering parsimony—such a low chi-square value with so many degrees of freedom cannot be interpreted in a straightforward manner. Nonetheless, this model fit shows that our hypothesis about the importance of bridge symptoms in explaining comorbidity is not a priori wrong.

The network approach is based on the hypothesis that symptoms are related *directly*. It is important to qualify this terminology to prevent misunderstandings. We intend the term "directly" to mean that relations between symptoms are real; that is, not spurious in the sense that a latent variable model assumes them to be. However, this does not imply that intermediate processes or attributes are not involved in these symptom-symptom relations. For instance, the influence of one symptom on another is likely to be mediated by, or instantiated in, a chain of processes that are not directly observable. Even the

[2] We did not collapse the six symptoms that overlap between MDD and GAD into three bridge symptoms because the log odds ratios between each pair of overlapping symptoms were not high enough to warrant such a collapse. A probable explanation for this is that some people, for instance, did report concentration problems in the depression section, but were unable to report those same problems in the generalized anxiety section because that section was skipped (e.g., because the respondent did not experience chronic anxiety).

influence of the symptom 'sleep disturbances' on 'fatigue', mundane as it may seem, will invoke various intermediate mechanisms concerning the homeostatic processes involved in sleep regulation (Achermann, 2004; Borbély & Achermann, 1999; Finelli, Baumann, Borbély, & Achermann, 2000). Thus, within a network framework, it makes perfect sense—and is naturally necessary—to introduce non-symptom causal processes such as homeostasis that partly explain relations between symptoms. Also, such processes may involve pathways that contain some of the other symptoms in the network; for instance, a lack of sleep may lead to a loss of concentration *via* fatigue. Finally, the causal effect of a symptom may feed back into that same symptom via a loop. For instance, fatigue may lead to a lack of concentration, which may lead to thoughts of inferiority and worry, which may in turn lead to sleepless nights, thereby reinforcing fatigue. In such a case, we have a vicious circle, or negative spiral, a well-known phenomenon to any practicing clinical psychologist. In some disorders, the existence of feedback loops is in fact considered to be a core aspect of the disorder; an example is panic disorder, in which 'fear of fear' appears to play a crucial role; for instance, when the fear of having a panic attack itself contributes to the occurrence of such an attack (McNally, 1994). It is therefore notable, and problematic, that in standard psychological measurement models, such phenomena cannot arise because latent variable models, being instantiations of a common cause structure, are *directed* graphs which, by definition, do not contain feedback relations[3] (Pearl, 2000).

Moreover, targeting such relationships between symptoms or processes that influence such relationships is a major goal of many successful therapeutic interventions such as *cognitive therapy* (e.g., lessen the impact of cognitions on relationships between symptoms: "If I do not finish all tasks I set out to do during the day, I am a worthless person and it is better for everyone if I were gone"; see Beck, Rush, Shaw, & Emery, 1979) and *exposure therapy* (i.e., breaking the link between seeing a particular object and responding to it with fear by repeatedly exposing a patient to the feared object; see e.g., Kamphuis & Telch, 2000; Rothbaum & Schwartz, 2002). It is therefore also problematic that such successful and common therapeutic interventions do not naturally arise from a latent variable perspective. This is not to say that targeting relations between symptoms is prohibited by a latent variable perspective; the more logical consequence of adopting such a perspective just seems to be to target the latent variable: eliminating the common cause will result in the disappearance of its indicators (i.e., the symptoms). In the case of major depression, for example, finding the common cause was therefore a major goal in research, with serotonin shortage being the most likely candidate. However, treatment with antidepressants that specifically target that shortage turned out to be beneficial for only some people, thereby ruling out serotonin as the common cause of depression symptoms (e.g., see Nierenberg et al., 2008). No other plausible common causes have ever been found, in our opinion due to the fact that there simply is no common cause that explains the entirety of depression symptoms.

[3] It is prudent to note that feedback loops can create considerable methodological difficulties in model fitting, because they lead to models that cannot be recursively estimated. However, given our present state of ignorance concerning the nature of comorbidity, we think it is more useful to construct a theoretical representation that is likely to be faithful to reality, than it is to construct a model based on a list of desirable computational properties.

An integrative method to visualize symptom associations through graphical models

Many of the effort in complex systems theory have been aimed at providing adequate visual representations of networks, and this has yielded a number of algorithms to optimally represent networks (Berg, Cheong, Kreveld, & Overmars, 2008; DiBattista, Eades, Tamassia, & Tollis, 1994; Herman, 2000), as well as freely available software to visualize them; most notable, in this respect, are the programs Cytoscape (Shannon et al., 2003 —used in constructing the graphs for this chapter), aiSee (http://www.aisee.com), and igraph (Csárdi & Nepusz, 2006 —used in this chapter for the detection of community structures). We therefore propose that the study of comorbidity through network models may best start by constructing insightful visualizations.

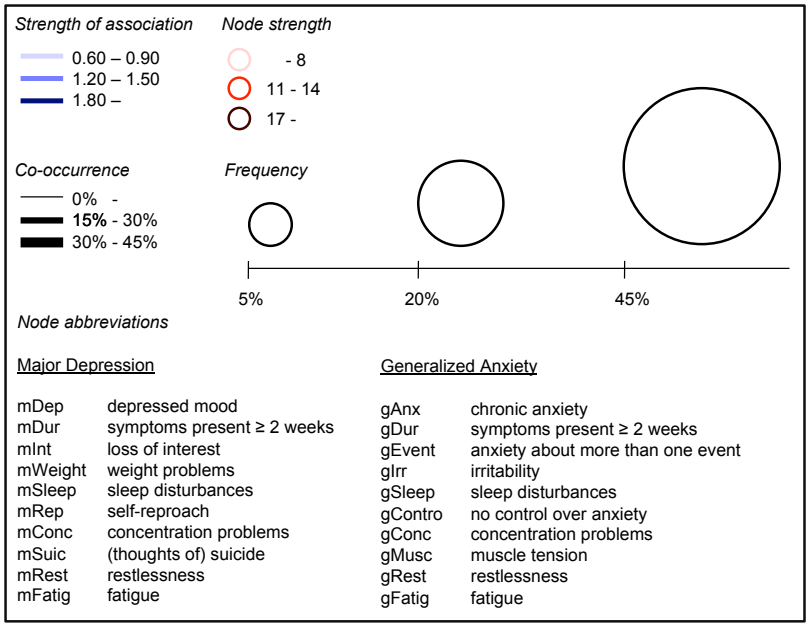

Figure 2.3: The key for the comorbidity networks shown in Figures 2.4, 2.5 and 2.6.

Among a plethora of possibilities to define and visualize both nodes and edges (see, e.g., Boccaletti et al., 2006; Krichel & Bakkalbasi, 2006), we propose an integrative method that, in our view, optimally visualizes key aspects of comorbidity on a symptom level. Figure 2.3 provides the complete key to such a *comorbidity network* for MDD and GAD, which is presented in Figure 2.4[4] First, the *thickness of the edges* is determined by the co-occurrence of two symptoms: the more two symptoms co-occur, the thicker the edge between them. Second, the *color of the edges* is determined by the

[4] This network is based on the NCS-R questionnaire that mostly contains dichotomous items. However, some of the items were ordinal or continuous (e.g., "How many pounds have you gained?"), and we dichotomized those according to the DSM-IV diagnostic algorithms. Details of the dichotomization process are provided at: http://www.aojcramer.com.

log odds ratio between two symptoms[5] (i.e., *strength of the association*; results available at: http://www.aojcramer.com): the higher the log odds ratio, the darker blue the edge between symptoms (please note that other options exist to define some measure of the strength of the association between two symptoms: for instance, tetrachoric correlations[6]). Third, the *size of the nodes* is determined by the raw frequency: the more frequent a symptom, the larger the node. Finally, the *color of the nodes* is determined by their individual *node strength* (see, e.g., Boccaletti et al., 2006; Krichel & Bakkalbasi, 2006). The node strength is simply the sum of the weights of all edges that are incident in that node. In the complex networks literature, the node strength is taken to be a measure of the *centrality* of a node such that the more strength, the more central a node is in the network.

In addition, we propose the following two rules for the positioning of the nodes in a comorbidity network (see also Figure 2.4): First, we propose that from left to right (i.e., the x-axis), non-overlapping symptoms of two disorders are placed in the middle of the graph. As such, one can immediately see whether comorbidity between two disorders runs mostly through the overlapping symptoms or (also) exists independently from them. Second, we propose that from top to bottom (i.e., the y-axis) the nodes are placed based on descending node strength. As such, one can immediately see which symptoms are more central in the network (i.e., top of the graph).

The basic structure of the depression and generalized anxiety comorbidity network

A few characteristics of the MDD and GAD comorbidity network stand out in particular (see Figure 2.4[7]). First, GAD symptoms are more frequent than MDD symptoms (i.e., GAD nodes are generally larger than MDD nodes). At first sight, this may appear at odds with the higher prevalence of MDD compared to GAD that is usually reported (Carter, Wittchen, Pfister, & Kessler, 2001; Kessler, Chiu, et al., 2005). However, on a diagnosis level, only respondents who display *a certain number* of MDD or GAD symptoms with *a certain duration* qualify for a diagnosis. Additionally, because of a hierarchical exclusion rule, the GAD diagnosis will not be assigned if its symptoms occur exclusively within the course of MDD (T. A. Brown & Barlow, 1992, 2001; L. A. Clark, Watson, & Reynolds,

[5] The *odds ratio* is the ratio of the odds of an event (e.g., suffering from loss of interest) occurring in one group (e.g., people who suffer from depressed mood) to the odds of that event occurring in another group (e.g., people not suffering from depressed mood). For cell counts in a 2x2 contingency table, the sample odds ratio equals $n_{11}n_{22}/n_{12}n_{21}$ (see Agresti, 2002). Since the odds ratio scales between zero and infinity, with a value of 1 signifying the absence of association, the odds ratio is not optimal for visualization in our network; therefore, we used the natural logarithm of the odds ratio. A log odds ratio of 0 (i.e., an odds ratio of 1) indicates that the event is equally likely in both groups. Please note that a high co-occurrence ($= n_{11}$) does not necessarily imply a high odds ratio. For example, (1) a high co-occurrence ($n_{11} = 500$), (2) almost no people who do not have both symptoms ($n_{22} = 3$), and (3) thus relatively many people who have one or the other symptom ($n_{12} = 15$ and $n_{21} = 100$) yields an odds ratio of 1 (500*3/100*15), signaling no association between those symptoms. Thus, co-occurrences and odds ratios show different aspects of a data set.

[6] In fact, we also computed tetrachoric correlations for the MDD and GAD symptoms with a full information maximum likelihood approach through which we dealt with the missing values that were Missing At Random (MAR). We found that the ordering of the symptoms in terms of their node strength was nearly the same as with log odds ratios.

[7] We have checked the stability of the results depicted in this figure by randomly splitting the sample in two and running all analyses for both groups separately. Those separate analyses revealed the same results and, therefore, we consider the components of Figure 2.4 to be stable.

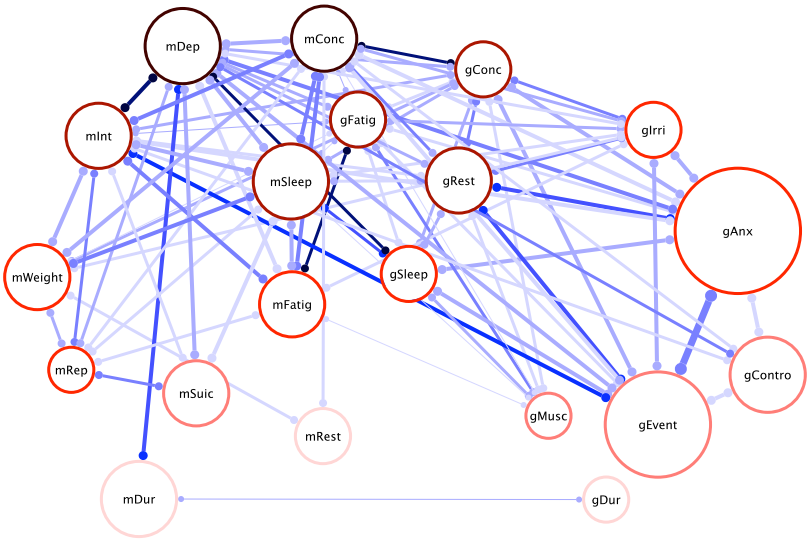

Figure 2.4: A comorbidity network for major depressive disorder (MDD) and generalized anxiety disorder (GAD). Larger nodes represent more frequent symptoms, darker circumference represents higher centrality, thicker edges represent higher co-occurrence, and darker edges represent stronger associations. Only edges with a log odds ratio higher than (+ or -)0.60 are represented. Centrally positioned nodes represent overlapping symptoms. Non-overlapping MDD symptoms are displayed on the *left* of the figure, and non-overlapping GAD symptoms on the *right*.

1995; Mineka, Watson, & Clark, 1998; Watson, 2005). Since MDD and GAD are highly comorbid (see, e.g., T. A. Brown et al., 2001, 1998; Mineka et al., 1998), such exclusion rules lower the prevalence of GAD artificially. Here, we consider data of all respondents who completed the MDD and GAD interview sections, regardless of whether or not they obtained diagnoses. As such, the network demonstrates that, when considering both subthreshold and threshold depression and generalized anxiety, symptoms of generalized anxiety are in fact more prevalent.

Second, if MDD and GAD are separate entities, we would have expected the edges to be thickest between symptoms of the same disorder (i.e., high co-occurrence). However, it is apparent that this is not the case in the network: Some of the thickest edges connect MDD with GAD symptoms; for instance, the thick edge between loss of interest (*mInt*) and reporting more than one event one worries about (*gEvent*). Also, we would have expected edges to be darkest blue between symptoms of the same disorder (i.e., high log odds ratios), but that is also not evident when inspecting the figure. In other words, associations between symptoms of one disorder are not stronger than between symptoms of different disorders. These findings are in line with an earlier hypothesis that MDD and GAD are hard to distinguish on a genetic level (Mineka et al., 1998) and, as such, raise the question of whether MDD and/or GAD are truly distinct disorders. We will return to this matter in more detail in the paragraph about the non-uniformity of diagnostic criteria.

Third, duration (*mDur* and *gDur*) is hardly associated with any of the other MDD

and GAD symptoms[8] (i.e., few edges are incident in those nodes). This may appear surprising since, in clinical practice, duration is key in determining the presence or absence of a mental disorder. However, if we consider medical illnesses as an analogy, the finding is potentially less surprising: Cancer will be diagnosed regardless of how long its symptoms (e.g., coughing up blood in the case of lung cancer) have been present.

Finally, the strongest evidence for comorbidity stems from strong associations that involve at least one overlapping symptom (e.g., between depressed mood, *mDep*, and sleep disturbances, *gSleep*). This apparent nontrivial role of overlapping symptoms in comorbidity stands in stark contrast to earlier findings regarding MDD, GAD, and other mental disorders (e.g., see Biederman, Faraone, Mick, & Lelon, 1995; Bleich, Koslowsky, Dolev, & Lerer, 1997; L. A. Clark & Watson, 1991; C. L. Franklin & Zimmerman, 2001; Kessler, DuPont, Berglund, & Wittchen, 1999; Seligman & Ollendick, 1998; Watson et al., 1995). We will return to this issue in more detail in the paragraph about overlapping symptoms.

It is crucial to note that the network is not necessarily complete. That is, this comorbidity network is based on the symptoms of major depression and generalized anxiety, but, naturally, it stands to reason to hypothesize the presence of factors—other nodes—that selectively influence some of the symptoms and are thus part of the network. For instance, it is well known that major life events, such as the loss of a loved one, can trigger major depression and, more specifically, there is evidence for selective influence of such personal tragedies on the more psychological symptoms of depression (e.g., depressed mood, thoughts of suicide; David, Ceschi, Billieux, & van der Linden, 2008; Kessler, 1997; Monroe, Harkness, Simons, & Thase, 2001). Also, there is evidence that traits such as neuroticism (mediated by rumination on sadness) and behavioral inhibition (i.e., shy, fearful, and withdrawn) can trigger the onset of depression and/or anxiety symptoms (e.g., see Hirshfeld et al., 1992; McNiel & Fleeson, 2006; Roelofs, Huibers, Peeters, & Arntz, 2008, 2008). Because such and other more "etiological nodes" are missing from this network, they are in a sense latent. However, such latent etiological nodes do not turn the MDD and GAD comorbidity network into a latent variable model: A network with multiple latent nodes that selectively influence some of the symptom nodes is not the same as a latent variable model in which one latent factor influences *all* and thus entirely explains relations between symptom nodes. Moreover, an unobserved variable is indeed latent, but not every unobserved variable automatically qualifies as a latent variable in the psychometric sense in which such variables are portrayed in latent variable models commonly used in data analysis.

The inequality of symptoms and its consequences for diagnostic cut-offs and the definition of a mental disorder

The focus in comorbidity research is on diagnoses, which means that inferences regarding comorbidity rest on summed scores that are obtained by counting symptoms. In latent variable modeling, such an unweighted summed score is either a sufficient statistic for

[8] The fact that duration is weakly associated with the other MDD and GAD symptoms cannot be explained by a skip structure that only allowed respondents to progress to the other symptoms' section if they fulfilled the duration criteria for depressed mood/loss of interest (MDD: more than 2 weeks) and chronic anxiety (GAD: more than 6 months): respondents with depressed mood/loss of interest for at least 3 days for more than 1 hour per day (MDD) as well as respondents with chronic anxiety for at least one month were allowed into the sections about the other symptoms.

the latent variable (e.g., see Andersen, 1973; Masters & Wright, 1984) or has a monotone likelihood ratio with that latent variable (Grayson, 1988). In both of these cases, inferences based on the summed symptom scores will often generalize to the latent variable. The unweighted summation of symptom scores implies that all symptoms are considered equal. Although thus formally consistent with latent variable modeling (Grayson, 1988), this assumption is highly problematic and may be the origin of some significant problems in comorbidity research. In a network approach, symptoms are likely to be actually unequal in terms of their *centrality*, a property that is not reflected in any latent variable model, and this has consequences for the comparability of equal summed scores.

Suppose that Alice displays two MDD symptoms—depressed mood and loss of interest—while Bob displays two other MDD symptoms —psychomotor and weight problems. On an intuitive level, it is plausible that Alice's symptoms are more likely than Bob's to eventually result in a full-fledged depression. In other words, some symptoms appear to be more *central* features of depression than others. The comorbidity network sustains this intuition. When considering the node strengths in Figure 2.4 (i.e., colors of the nodes), one immediately sees that, indeed, depressed mood (*mDep*) and loss of interest (*mInt*) are far more central in the network than are psychomotor (*mRest*) and weight problems (*mWeight*). In other words, the same summed score of Alice and Bob may not adequately capture that the symptoms of Alice result in a higher probability of developing other MDD symptoms—and thus augment the probability of eventually developing depression—compared to Bob's symptoms. Hence, summed scores appear to be *incomparable*, at least with respect to elucidating which people with subthreshold depression problems are at more risk of developing MDD. Naturally, such symptom inequalities are widely recognized among psychiatrists and clinical psychologists, and they do occasionally appear in DSM-IV (e.g., depressed mood and loss of interest as central features of major depression); the problem is, however, that the models that underlie current comorbidity research do not naturally allow for them.

If our line of reasoning is correct, and there is no latent variable that screens off correlations between symptoms (*a latent variable model renders all symptoms equally central and exchangeable*[9]), then the inequality of symptoms in terms of their centrality also renders diagnostic cut-offs open to debate. We are certainly not the first ones to point out that diagnostic cut-offs appear to be arbitrary (e.g., see Gotlib, Lewinsohn, & Seeley, 1995; Lilienfeld & Marino, 1999; Maier, Gänsicke, & Weiffenbach, 1997; Solomon, Haaga, & Arnow, 2001). For instance, there are individuals who do not meet diagnostic criteria for MDD yet appear to be psychosocially as dysfunctional as individuals who are diagnosed with MDD; that is, the consequences of subthreshold MDD problems may not always be distinguishable from those of diagnosed MDD. With the network approach, we offer a potential explanation of such findings. Suppose that Alice displays four MDD symptoms and Bob five. The diagnostic cut-off of criterion B for MDD is five, so Alice would not be diagnosed with MDD while Bob would. So far so good, but now suppose that Alice's symptoms are all highly central in the MDD network while Bob's are more peripheral. Is it, in such a scenario, plausible to conclude that Bob is depressed and Alice is not depressed? In other words, based on diagnostic cut-offs, we may fail to disentangle symptom-specific effects, because such cut-offs do not take into account the centrality of symptoms.

This brings us to another important point: namely, the definition of a mental disorder, generally conceptualized as "Disorder A is X or more symptoms out of Y possible

[9] It is important to note here that within a latent variable framework, factor loadings cannot be measures of symptom centrality as we view the concept, since those loadings are simply reliability estimates: the higher the factor loading, the more reliably an indicator "represents" the common cause.

symptoms". According to a latent variable perspective, it is not only perfectly defensible to entertain such a definition, but the definition is the same for every single individual; that is why Alice is not depressed and Bob is. However, if symptoms are not exchangeable in terms of their centrality, as we think is plausible, one cannot help but question such a definition of a mental disorder. In other words, if diagnostic cut-offs alone are no longer the demarcation line above which someone suffers from a particular mental disorder, then how do we define a mental disorder?

From a network perspective, there are several possibilities to define what constitutes a mental disorder. As a starting point, we propose to define a disorder as a cluster, a set of nodes (symptoms) that are strongly connected. Now, from a graph theoretic perspective, there are multiple ways to define in what sense a set of nodes is strongly connected (see, e.g., L. J. Hubert, 1994). First, let us call the giant network consisting of all symptoms of all mental disorders (i.e., the entire *symptom space*) as defined in the DSM-IV, graph G. Then a subgraph H (for instance, consisting of all MDD symptoms) is a cluster of G if and only if the minimum node strength of H is larger than the minimum node strength of H + n, with n being any other node adjacent to H (*Definition 1*). It is also possible to define a subgraph H as a cluster of G if and only if the minimum of the average distance between all nodes in H is strictly smaller than that of H + n for any node n in G (i.e., closeness; see, e.g., Boccaletti et al., 2006) (*Definition 2*). Other definitions are possible, and it is—in our opinion—up to future debate and research to determine which is the most sensible one. Second, now that we have hypothetically defined the cluster of all possible symptoms of a disorder, we need to determine when such a cluster is disordered. One plausible candidate is a modified version of the diagnostic cut-off; for example, in the case of MDD, at least three *of the most central* symptoms in the entire MDD cluster (with "central" either defined as the nodes with the largest node strengths, or as the smallest average distance within the cluster). In contrast to a latent variable perspective, both definitions acknowledge the centrality differences of symptoms but, at the same time, accept the inevitable fact that some form of a diagnostic cut-off is needed to disentangle people with and without a disorder.

A related point concerns the *external* effects of different symptoms. One readily imagines extending a network with variables that are not part of the disorder itself, but constitute nontrivial consequences of many mental disorders (e.g., losing one's job, lowered educational achievement, or suicide attempts). It is interesting to note that, under the assumption of a latent variable model, it is the latent variable that has a direct relationship with external effects, and not the symptoms. Due to the absence of a direct relationship between a symptom and an external effect, this means that a symptom can never be statistically independent of such an external effect, given another symptom. Thus, for instance, a suicide attempt by someone with thoughts of suicide and concentration problems (and three other symptoms resulting in a diagnosis of major depression) is entirely attributable to the overarching latent depression and, given the thoughts of suicide, the concentration problems are thus still associated with the suicide attempt. In our view, it would be more logical to hypothesize a direct relationship between thoughts of suicide and a suicide attempt and a weaker or perhaps even nonexistent relationship between concentration problems and a suicide attempt. In the same vein, it appears to make sense to envision a stronger relationship between concentration problems and losing one's job than between losing weight and losing one's job. This differential impact of symptoms on external effects is not possible in a latent variable model, whereas it is very easily envisioned within a network perspective.

Centrality differences between symptoms imply that there probably will be pathways to comorbidity that are more likely (i.e,. strong connections between symptoms that are central in the network) than others. Figure 2.4 confirms this idea: One likely pathway

to comorbidity connects depressed mood (*mDep*) with sleep problems (*gSleep*) and anxiety (*gAnx*). Less likely pathways involve psychomotor problems (*mRest*) because this symptom has such weak associations with the other symptoms in the network. Naturally, inspecting a graph is not enough to draw any solid conclusions on the pathways to comorbidity between MDD and GAD, but we do think it is evident that the network approach could contribute to finding answers to this question, if only because the visual representation of a network immediately leads to a host of interesting hypotheses.

Non-uniformity of mental disorders

Quite a few scholars are *essentialists* in describing the relationship between the two main diagnostic categories "disorder" and "no disorder" that are based on diagnostic criteria and the real world (e.g., see Haslam, 2000, 2002; Lilienfeld & Marino, 1999): The diagnostic criteria we use result in a distinction between disordered and non-disordered people that also exists in the real world. Seductive as this line of reasoning may seem, in order for it to be true, two conditions must be satisfied. First, a mental disorder must have *defining features* such that everyone, based on those defining features, could be assigned to the "disorder" category (i.e., defining features are present) or the "no disorder" category (i.e., defining features are absent) provided that these features were known with certainty. Second, as a result, all members of the same category must essentially be the same with respect to those defining features (i.e., *uniformity*). Down's syndrome is a good example of a medical disorder that satisfies those two conditions: The syndrome has one defining feature, the presence of all or part of an extra 21^{st} chromosome, and everyone with Down's syndrome possesses that defining feature while everyone without Down's syndrome does not possess it.

This line of reasoning is unlikely to hold for mental disorders. First, quite a few mental disorders do not have defining features, at least not in an essentialist sense. For example, besides depressed mood or loss of interest, which must always be present for a person to be diagnosed as having MDD, *any* constellation of five symptoms (i.e., features) will suffice to fulfill criterion B for MDD. When any such constellation of symptoms is present for at least two weeks in an individual, then that individual will be assigned to the "MDD" category, otherwise to the "no MDD" category. This renders the core features of depression non-defining because, for instance, someone with the feature "depressed mood" could end up in de the "MDD" category—because he or she suffers from five or more symptoms for more than two weeks—as well as the "no MDD" category because he or she suffers from less than five symptoms or the symptoms are present for less than two weeks. Second, as a result of the lack of truly defining features, the "basket" with depressed people does not contain uniform members: Pete is depressed because he suffers from sleep disturbances, fatigue, concentration problems, depressed mood, and psychomotor problems; while Anne is depressed because she suffers from depressed mood, loss of interest, self-reproach, weight problems, and thoughts of suicide.

As such, one must wonder whether the distinction between "disorder" and "no disorder", as we have defined it in our diagnostic criteria, actually exists in the real world. Latent variable modeling schemes posit the existence of such a categorical system (in a latent class model) or a continuous one (in a factor or item response theory [IRT] model) as a hypothesis. Hence, such models are consistent with the hypothesis that we may one day find out "what depression really is"; that is, latent variables may "become" observed through of a refinement of the conceptual and measurement apparatus used to study them (e.g., Bollen, 2002; Borsboom, 2008b). However, in the absence of such refinements, the acceptance of the latent variable hypothesis depends at least partly on its explanatory values (Haig, 2005), and in the context of comorbidity research these

explanatory virtues are, at present, quite limited. That is, apart from the fact that such a model would explain why correlations between symptoms are positive and that it more or less fits the observed frequency of symptom patterns, there is little that speaks in its favor.

When studying comorbidity based on diagnoses, this inevitably leads to the question of what we actually observe when two disorders covary: genuine covariation between two real disorders, or covariation between certain constellations of symptoms we have designated to be disorders, but that are in fact not indicators of the same latent variable? This issue, of course, has generated a heated debate throughout the history of psychiatry and clinical psychology (Haslam, 2000, 2002; Jablensky, 2007; R. E. Kendell, 1975; D. F. Klein, 1978; Krueger & Markon, 2006; Lilienfeld & Marino, 1999; Richters & Hinshaw, 1999; Spitzer, 1973, 1999; Spitzer & Endicott, 1978; Wakefield, 1992, 1999a, 1999b; Zachar, 2000; Zachar & Kendler, 2007). The network approach could contribute to finding an answer to this question in two ways: first, by utilizing techniques to find what is called a *community structure*, and second, by reconceptualizing the question itself, and thereby the range of possible answers.

The community structure of a network refers to the existence of at least two clusters of nodes, such that the nodes within a cluster are highly connected with one another, but only modestly or sparsely with the nodes within another cluster (see Newman, 2006; Newman & Girvan, 2004). We analyzed the community structure of the MDD and GAD comorbidity network twice with a spinglass algorithm (for technical details, see Reichardt & Bornholdt, 2006): one time with co-occurrence between symptoms as edge weights and one time with the log odds ratios between symptoms as edge weights. The results are in line with the notion that there is no essential distinction between MDD and GAD, as has also been found in behavioral genetics and diagnostics research (Mineka et al., 1998; Wadsworth, Hudziak, Heath, & Achenbach, 2001): Our network reveals no community structure whatsoever, regardless of which edge weights were used; that is, the comorbidity network did not differ from a random network in terms of connectivity between nodes. These results suggest that MDD and GAD may not be separate entities. Naturally, this conclusion may be different for other mental disorders.

We are by no means pioneers when claiming that boundaries between diagnostic categories are *fuzzy*, for this phenomenon was noticed quite some time ago (e.g., see R. E. Kendell, 1975; D. F. Klein, 1978; Spitzer, 1973; Spitzer & Endicott, 1978). However, earlier ponderings have not included an account of *why* the boundaries are fuzzy and, in our view, a network approach offers such an explanation. If we are indeed correct to assume that a mental disorder is best conceptualized as a network of symptoms and—consequently—comorbidity is best viewed as a network of symptoms of two disorders, then boundaries are fuzzy *because they simply do not exist*. And the reason that they do not exist lies in the fact that the networks are not isolated from each other. The very fact that there are bridge symptoms precludes such a situation from occurring. As a result, we can draw the line between disorders A and B everywhere in the network. For instance, we could draw a boundary between MDD and GAD such that MDD contains only non-overlapping MDD symptoms while GAD contains its own symptoms and the overlapping MDD symptoms. Or, we could draw a boundary such that MDD only contains non-overlapping MDD symptoms and GAD only its non-overlapping symptoms. In other words, from a network perspective, the DSM-IV-defined boundary between MDD and GAD is no more defensible than any other boundary.

The network perspective offers an intermediate position between essentialism and conventionalism regarding mental disorders and the comorbidity that exists between them. On the one hand, there is a sense in which the delineations of mental disorders are arbitrary (there is no preferred line that separates the relevant networks). On the other

hand, since realizations of common causes for symptom clusters cannot be detected, the actual phenomenon of comorbidity is not a matter of convention, since it depends on causal patterns that exist in the real world independent of the researcher who studies them. Although mental disorders can be defined as a network in various ways, which may reflect mainly pragmatic concerns, comorbidity will remain regardless of how one draws the lines. In this sense, comorbidity may be more real than the mental disorders on which it is defined.

This is consistent with, and may offer an explanation of, results typically found in quantitative behavior genetics. Through twin studies and related methodologies, it has been established that a considerable portion of the individual differences in anxiety and depression, as well as many other psychological variables, is determined by genetic factors (Boomsma, Busjahn, & Peltonen, 2002; Kendler, Gardner, Neale, & Prescott, 2001; McGue & Christensen, 2003). Much research has focused on determining the genes responsible for this fact, but so far these efforts have been moderately successful at best, with the typical result being that individual polymorphisms do not account for more than a minor portion of the phenotypic variance (e.g., 1% or 2% at best). Thus, such phenotypes are highly polygenetic. The network account explains this naturally: It is likely that the strength of connections between symptoms (e.g., the relation between lack of sleep and irritability) differs over individuals, and it is also likely that these individual differences are at least partly under genetic control. However, a network of k nodes consists of k^2 - k relations between distinct nodes (380 possible relations for the network in Figure 2.4), and it is rather unlikely that the strength of *each* of these relations stands under control of the *same* genes. Thus, the network approach is not only consistent with the fact that most psychological phenotypes are polygenetic, but may actually offer an explanation of that fact. In addition, the approach suggests that gene-hunting efforts may be better served by relating polymorphisms to the *relations* between symptoms, rather than to *composites* of symptoms such as total scores on questionnaires.

The possibility of individual differences in a network structure raises the question of whether a uniform definition of comorbidity exists. For example, is there a particular *sequence* in which two comorbid disorders arise that holds for every single individual? At first sight, this appears to be unlikely. However, even though there may be individual differences in qualitative structure and quantitative characteristics of networks, statistical considerations regarding the average strength of connections may suggest pathways that are more or less prevalent in the population.

For instance, in contrast to Moffitt et al. (2007), who found that MDD and GAD were equally likely to be the first in the comorbidity sequence, the MDD and GAD comorbidity network (see Figure 2.4) does suggest the existence of a general pathway: namely, from MDD to GAD. First, because the non-overlapping MDD symptoms are not highly associated with one another, it does not appear to be very likely that someone with a few non-overlapping MDD symptoms will progress to other non-overlapping MDD symptoms. Second, a pathway from non-overlapping to overlapping MDD symptoms to GAD symptoms could be more likely because of stronger associations between those types of symptoms. The converse scenario—that is, from GAD to MDD—appears to be less likely in this particular network. In general, associations between non-overlapping GAD symptoms are relatively strong, at least stronger than between the symptoms of MDD, and, most importantly, more or less as strong as associations between non-overlapping and overlapping GAD symptoms. As such, when in the GAD network, to progress quickly from a few non-overlapping GAD symptoms to overlapping GAD symptoms and from there to MDD symptoms, does not appear to be more likely. Instead, it appears to be equally likely that someone stays in the GAD network without progressing to MDD symptoms. Given the structure of this particular MDD-GAD network, we therefore

hypothesize that Neale and Kendler (1995) are correct in concluding that the most likely pathway could indeed be from MDD to GAD.

Naturally, further research involving the time course and etiology of mental disorders is required to test this hypothesis. It should be noted, however, that the hypothesis follows naturally from a (tentative) causal interpretation of the network: the stronger the association between symptoms, the more likely that one symptom will lead to another. Furthermore, a causal explanation of a network suggests that some symptoms within a disorder put one at greater risk for comorbidity than do others. To the contrary, one does not get these implications from either unidimensional or two-dimensional latent variable models that assume exchangeable symptoms, save for measurement precision (see Bollen, 1989 for a good explication of this point). Thus, studying the etiology of symptoms may offer interesting insights with respect to the question of whether symptom development is best conceptualized in terms of a latent variable model, or in terms of a network perspective. We therefore consider the direction of research efforts toward the study of temporal dynamics of symptoms to be essential.

Symptom overlap between disorders

A final problem with current comorbidity research has to do with the fact that many disorders share a number of symptoms: sleep disturbances, fatigue, restlessness, and concentration problems in the case of MDD and GAD (APA, 1994). The obvious problem of such symptom overlap is that it raises doubt as to whether comorbidity is a real phenomenon: If we would remove overlapping symptoms from our diagnostic system, would comorbidity estimates look more or less the same, or is it that comorbidity is just that, symptom overlap? The latter does not appear to be true. Numerous researchers have approached this problem via different angles and with respect to different disorders, and the majority have reached the same conclusion: Yes, there is considerable symptom overlap between some disorders, but it seems highly unlikely that this overlap explains most systematic covariation between those disorders (Biederman et al., 1995; Bleich et al., 1997; C. L. Franklin & Zimmerman, 2001; Kessler et al., 1999; Seligman & Ollendick, 1998).

However, there are reasons to argue that some of the methodological approaches to study the effects of symptom overlap are problematic, rendering the conclusions based on such approaches open to debate. For instance, Bleich et al. (1997) removed symptoms that overlapped between posttraumatic stress disorder (PTSD) and MDD and re-diagnosed Israeli combat veterans who were already diagnosed with PTSD and/or MDD. The results showed that, after the removal of the overlapping symptoms, 98% (95%) of the veterans with lifetime (current) MDD were re-diagnosed with MDD, whereas 70% (55%) of the veterans with lifetime (current) PTSD were re-diagnosed with PTSD. Besides the fact that the re-diagnosis percentage of both lifetime and current PTSD is somewhat low, the problem with this approach is that re-diagnosing someone with MDD without overlapping symptoms does not prove that symptom overlap does not play a role in the *etiology* of comorbidity between MDD and another disorder.

Suppose that someone endorses eight MDD symptoms, three of which overlap with GAD. Two problems arise here. First, the effect of removing the overlapping symptoms depends on the diagnostic cut-off: This person will be re-diagnosed with a cut-off of five while with a cut-off of four, there will be no re-diagnosis. Hence, conclusions about the effects of removing overlapping symptoms depend entirely on diagnostic cut-offs that, as we noted earlier, are at least partially arbitrary. Second, and more important, it is impossible to exclude that a re-diagnosis actually signals the major impact of overlapping symptoms in explaining the etiology of comorbidity: What if overlapping symptoms are

relay stations that trigger the onset of symptoms in the entire network, resulting in a comorbid diagnosis? As such, a subsequent re-diagnosis does not have to signal the relative unimportance of overlapping symptoms. To the contrary, it could be justifiable taken to mean that overlapping symptoms have a seminal role. They cause comorbidity with such a profound effect on the network that removing them does not affect the initial diagnosis: the damage has already been done.

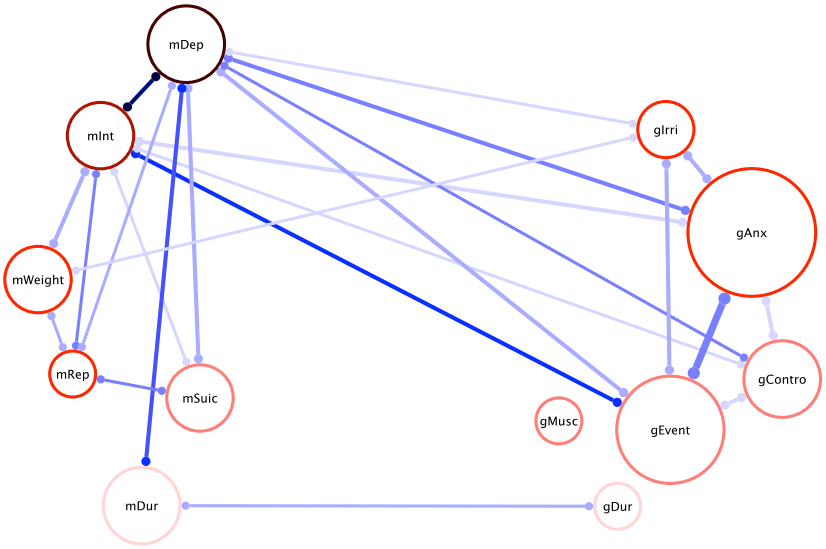

Figure 2.5: The comorbidity network for major depressive disorder (MDD) and generalized anxiety disorder (GAD) after removal of the overlapping symptoms and their bivariate associations with the other symptoms. This network is based on exactly the same four characteristics as the full network in Figure 2.4.

This is not to say that studying the effects of removing overlapping symptoms is a bad idea per se. We think it is a useful starting point, but (a) the effects of removing overlapping symptoms are perhaps better studied on a symptom level instead of on a diagnosis level, and (b) the matter should be investigated further; for instance, by not removing overlapping symptoms but by separately analyzing a subgroup: people who display one or more overlapping symptom pairs. Thus, we first investigate the impact of removing the six symptoms that overlap between MDD and GAD, as well as their associations with all other symptoms from the comorbidity network in Figure 2.4, resulting in Figure 2.5 (see Figure 2.3 for the key). This figure confirms our initial suspicions: without the overlapping symptoms, not much comorbidity seems to remain. In fact, only depressed mood (*mDep*) and loss of interest (*mInt*) have some relatively strong connections with GAD symptoms such as anxiety (*gAnx*), loss of control (*gContro*), and number of events that cause worry (*gEvent*).

Next, we performed the subgroup analysis: We thus computed log odds ratios, co-occurrences, frequencies, and node strengths for only those respondents who displayed at least one *pair* of overlapping symptoms (e.g., both MDD and GAD concentration

problems; $N = 1059$)[10]. Figure 2.6 presents their comorbidity network without the overlapping symptoms (see Figure 2.3 for the key). This figure leaves no room for doubt about the importance of overlapping symptoms: All symptoms are more frequent and co-occur more frequently, and having one symptom increases the odds of having another substantially (and thus the node strength) compared to the comorbidity network in Figure 2.5. Taking all results together, it is likely that overlapping symptoms play a more important role in explaining comorbidity than was originally thought.

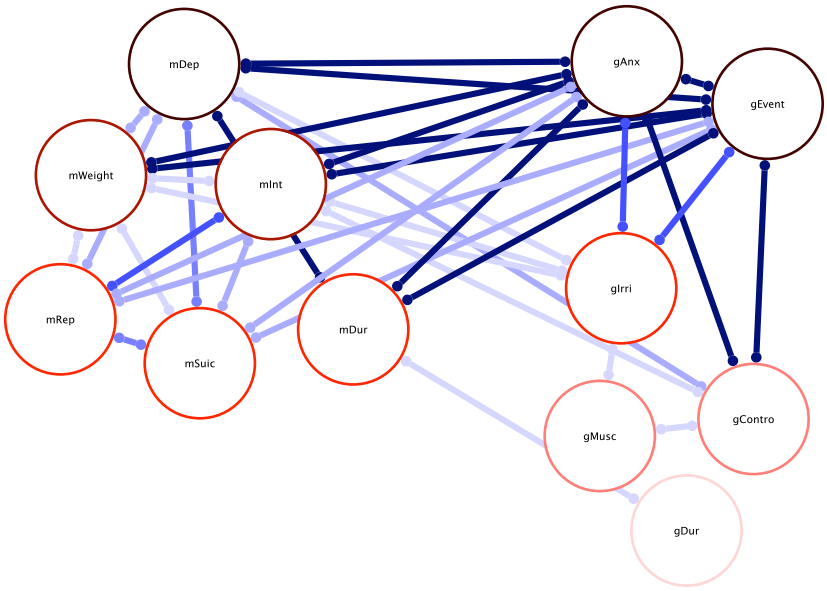

Figure 2.6: A comorbidity network for major depressive disorder (MDD) and generalized anxiety disorder (GAD) for those respondents ($N = 1059$) who displayed at least one pair of overlapping symptoms. This network is based on exactly the same four characteristics as the network in Figure 2.5.

Conclusions and future directions

In this chapter, we have introduced a radically different conceptualization of mental disorders and their symptoms: namely, the *network approach*. Under the assumption of such an approach, a mental disorder is a network of symptoms that stand in direct, possibly causal, relations to one another. Comorbidity between mental disorders is then conceptualized as direct relations between symptoms of multiple disorders. We have argued that such an approach bears a closer resemblance to the reality of mental disorders and comorbidity between them, as it allows for (1) multiple etiological processes that interact in causing symptoms, (2) interindividual differences in the manner in which a constellation of symptoms is contracted, (3) direct relations between overlapping symptoms, and

[10] The contingency tables, as well as the computational script (made in R), are available at: http://www.aojcramer.com. We have checked the stability of the results depicted in Figure 2.6 by randomly splitting the sample in two and have run all analyses for both groups separately. Those separate analyses revealed the same results, and therefore, we consider the components of Figure 2.6 to be stable.

(4) inequality of symptoms. Also, we have proposed an integrative method, based on bivariate associations, to visualize comorbidity networks.

Based on such an empirical network for major depression and generalized anxiety, we showed that a network approach results in a host of realistic and testable hypotheses that are not naturally accommodated by latent variable models. First, it is likely that there exist pathways to comorbidity through the symptom space that are more likely than others (e.g., via core psychological symptoms such as depressed mood and loss of interest). Second, it is plausible that those pathways generally follow the same direction (e.g., we found that comorbidity from major depression to generalized anxiety appeared to be more likely than the other way around). Finally, overlapping symptoms play a more than trivial role in explaining the roots of comorbidity (i.e., we showed that symptoms of major depression and generalized anxiety were more strongly connected in people who displayed at least one pair of overlapping symptoms).

The present work bears interesting relations to that of Van der Maas et al. (2006), who showed that the positive manifold of correlations between various IQ tasks—often thought to result from a single latent variable, general intelligence—may result from a dynamical system in which a network of bidirectionally related cognitive processes beneficially interact with one another during development (i.e., the *mutualism model*). The mutualism model serves as an excellent starting point for developing a unified theory for mental disorder networks because of their similarities. For instance, the mutualism model is a *dynamical system* (Alligood, Sauer, & Yorke, 1997) (for examples of dynamical systems in other areas of psychology, see Cervone, 2004; Shoda, LeeTiernan, & Mischel, 2002; van Geert, 1998). Such a system consists of a set of possible states with a rule that determines the present state in terms of past states. At any point in time, dynamical systems are in a particular state and that state can be represented as a point in *state space*. If a dynamical system evolves long enough, then it will encounter one or more *attractors* in state space: regions in state space that the system will move towards and enter. In state spaces with more than one attractor, some systems tend towards one attractor and remain there in a stable state (i.e., *monostable systems*; see, e.g., Pisarchik & Goswami, 2000).

The mutualism model is an example of such a monostable system. Like the mutualism model, mental disorders are also dynamical systems that evolve over time. However, unlike the mutualism model, mental disorder networks are probably minimally *bistable systems* with a "disorder" attractor state and a "no disorder" attractor state between which the system oscillates. For example, in a substantial number of people who suffer from major depression, it is a well-established fact that depressive symptoms come (i.e., the system moves towards a "depressed" attractor state), and go (i.e., the system moves towards a "not depressed" attractor state), either through therapeutic intervention or spontaneous remission (e.g., see Posternak & Miller, 2001). Some mental disorders may be *multistable systems* with the system oscillating between more than two attractor states. It is possible that bipolar II disorder is a system that oscillates between hypomania, major depressive episodes, and under the influence of therapeutic interventions, remission states. Dynamical systems theory can be used to predict the *trajectory* of a system in the state space; that is, future states of the system can be predicted from earlier states, a technique that is, for instance, widely employed in weather forecasting (e.g., see Palmer, 2001). Analogously, such techniques could in the future be used to predict trajectories of a variety of mental disorders, given the initial state of a network for an individual. If there are individual differences in the precise structure of networks, this may require person-specific network structures to be determined for each individual separately, as is, for instance, possible through the analysis of intra-individual time series (Hamaker, Nesselroade, & Molenaar, 2007; Molenaar, 2004).

The trajectory of any mental disorder as dynamical system cannot be adequately predicted without taking external variables into account. One important feature of many mental disorders is that all or most symptoms are positively correlated. As such, when modeling the reality of mental disorders from a dynamical systems perspective, if people enter the network by displaying one symptom, this symptom will quickly turn other symptoms "on". As a result, the trajectory of such a system will be predictable and unrealistic: everyone will "contract" the mental disorder. In reality, there are many external variables that mitigate relationships between symptoms: good news that prevents someone progressing from depressed mood to thoughts of suicide, homeostasis due to which someone with sleep difficulties will not stay fatigued indefinitely, and so on. Such external variables thus play a critical role in determining toward which attractor state the system moves, and, as such, must be included in mental disorder systems.

Also, we should take into account the possibility that the entire symptom space network displays characteristics of a *small world* (e.g., see Barrat & Weigt, 2000; Rubinov et al., 2009; Watts & Strogatz, 1998). A small-world network is a highly clustered network with relatively short characteristic path lengths (i.e., it takes relatively few steps to "travel" from one node in the network to another). Networks with such properties are frequently found, ranging from the power grid of the western United States through the neural network of the worm *Caenorhabditis elegans*. If a general mental disorder system would indeed also display small-world features, it potentially offers a powerful explanation of the generally high comorbidity between mental disorders (i.e., short characteristic path lengths). Also, it might reconfirm the existence of distinct symptom clusters that represent distinct mental disorders (i.e., high clustering).

Finally, any adequate general network model for mental disorders must encompass the fact that mental disorders as systems are essentially *complex* (e.g., see Cilliers, 1998): Because of the interplay between the individual components (i.e., symptoms) of the system and the interaction between the system and its environment, the system/disorder as a whole cannot be fully understood by analyzing its individual components. Also, these interactions change over time, and this can result in *emerging properties*, properties of the system that are not evident from inspecting the individual components. In complexity research, rapid advances are made with respect to modeling emerging properties in complex systems, and the network approach for mental disorders could benefit from those advances (see, e.g., Paik & Kumar, 2008; Solé, Ferrer-Cancho, Montoya, & Valverde, 2000). An important additional question is how dynamical properties of complex systems relate statistically and conceptually to interindividual differences as commonly analyzed with latent variable models (Molenaar, 2003).

As such, multiple insights from various research disciplines may be further developed and combined into a general psychometric theory of mental disorders as networks. Such a theory should, in our view, address the dynamical nature of causal systems (i.e., model that tracks the development of a mental disorder network over time), allow for representing the influence of external variables (e.g., treatment that potentially turns symptoms "off"), and allow for an adequate conceptualization of causal relations between symptoms. Advances in the areas of complexity and dynamical systems may be of considerable help in constructing such a theory. Also, given the relevance of results from various disciplines (e.g., mathematics, physics, and computer science), the construction of a viable psychometric theory based on these ideas is likely to involve the integration of theory and methods from different fields, and we therefore hope to attract the attention of scholars from a wide variety of disciplines. The need for a general theory of this type is, we think, evident: We have been looking at mental disorders through the wrong psychometric glasses, and it is high time for us to craft new ones.

Chapter 3

Complex realities need complex theories

Abstract

This chapter was written in response to a set of commentaries (abstracts of these commentaries available in Appendix A and full texts available at http://www.aojcramer.com), which were written in response to the previous chapter. The commentators comprised scholars from various disciplines, ranging from philosophy and clinical psychology to psychometrics. We thank these commentators for their suggestions and critiques that have aided sculpting the ideas that are presented in this chapter. Although critical at times, the majority of commentators agree on one thing: Our network approach might be the prime candidate for offering a new perspective on the origins of mental disorders. After briefly discussing the gist of the commentaries, we elaborate in our response on refinements (e.g., cognitive and genetic levels) and extensions (e.g., to Axis II disorders) of the network model, as well as discuss ways to test its validity.

Adapted from: Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Complex realities require complex theories: Refining and extending the network approach to mental disorders. *Behavioral and Brain Sciences, 33*, 178-193.

In Chapter 2, we have proposed a network view of mental disorders, in which systematic covariation between symptoms is explained by direct relations between the symptoms themselves. The approach breaks radically with the dominant doctrine, in which disorders are considered to be common causes of their symptoms (i.e., the *latent variable perspective*). We were pleased to see that many commentators view the *network approach* as a potential substantive theory of mental disorders. Given the varied set of responses, many of which proposed worthwhile empirical research avenues and theoretical extensions of the approach, we have fortunately succeeded in bringing together researchers from different fields to reconsider what disorders are and how we should investigate them.

One of the most surprising and noteworthy facts about the present set of commentaries concerns what they do *not* contain: Very few commentators attempt to defend the received view that underlies many current approaches to psychopathology: that is, the latent variable perspective. We take this to imply that the time is ripe for a change of perspective. In addition, the comments have strengthened our conviction that, with the necessary refinements and extensions, "inference to the best explanation" could ultimately lead us to the network approach as *the* substantive theory of mental disorders (Haig, 2009). Certainly, **Rothenberger, Banaschewski, Becker, and Roesner** argue that the network approach is complex with its "manifold interactions between symptoms", but we agree with them even more that this reflects reality. And as we will argue here, complex realities require complex theories.

In this response, we discuss the most important extensions, refinements, investigative tools, and objections voiced by the commentators according to the following themes. First, several commentators argued that network models can and necessarily must include latent variables (e.g., **Haig & Vertue; McFarland & Malta**). In Section 2, we explain why some relations qualify for such a measurement model—and are thus likely to be incorporated into a network model—while others do not (e.g., depression as common cause of a cluster of symptoms). Other commentators provided excellent suggestions for refinement of the network model in order to include genetic, neurological, and cognitive levels of explanation (e.g., **Rubinsten & Henik; Yordanova, Kolev, Kirov & Rothenberger**), which we discuss in Section 3. Additionally, in Section 4, we discuss ways to test the network model, as suggested by several commentators (e.g., **Davis & Plomin; Fleeson, Furr & Arnold; van der Sluis, Kan & Dolan**). Section 5 investigates the possibility of extending the network approach to other disorders (e.g., Axis II personality disorders: **Bornstein & Ross**). Section 6 focuses on an important question, posed by several commentators, as to what constitutes a mental disorder (**Haslam; Hood & Lovett; Zachar**). Finally, commentators raised methodological objections that were claimed either to invalidate the network model we suggested (e.g., **Danks, Fancsali, Glymour & Scheines; Krueger, DeYoung & Markon**), or to sustain a common cause view on mental disorders (e.g., **Belzung, de Villemeur, Lemoine & Camus; Humphry & McGrane**). In Section 7, we discuss these issues and argue that—despite methodological difficulties that have to be addressed in the future—the network model should be viewed as the prime candidate to elucidate the origins of mental disorders.

Latent variables in the network approach

Markus and **Molenaar** remark that, if the network approach is to move from a mere representation of the data to a possible representation of the underlying causal and functional relations between its components, one requires a way to deal with the fact that the observations (i.e., symptom reports) are likely to be imperfect indicators of these components (i.e., the actual symptoms). These commentators note that, if measurement error is neglected, relations between symptoms can be inaccurately represented because

of attenuation effects. The only way to deal with this is to invoke latent variables into the model. Other commentators express this concern as well when discussing symptoms that should be measured in multiple ways (**Krueger et al.; McFarland & Malta**) or non-symptom causal processes that mediate the direct relations between symptoms (**Belzung et al.; Danks et al.; Haig & Vertue; Humphry & McGrane**). Our response is simply to acknowledge that this is the case; in fact, in Chapter 2, we specifically hint at this idea in the last paragraph of Section 4.

We construct the situation as follows: At the level of individual symptoms, we take symptom reports to be measures. If measurement error is to be accounted for at this level, one would indeed need multiple indicators per symptom and a parallel extension of the network model with latent variables; for example, a network model for depression could include sleep disturbances as a latent variable measured with three observable indicators (i.e., clinical interview, polysomnography, laboratory observation; see **McFarland & Malta**). Figure 1 depicts such a network model with sleep disturbances and weight problems as latent variables. Also, a model in which some non-symptom causal processes are latent because they are measured in multiple ways (e.g., "major life events" for depression) is easy to conceive, and we welcome the development of such extensions of the model (**Belzung et al.; Danks et al.; Haig & Vertue; Humphry & McGrane**).

The central tenet of Chapter 2 is, therefore, not to shun latent variables completely. For example, a measurement model that includes a latent variable makes perfect sense in case of the symptom "insomnia" with three indicators. This is because (1) a natural referent exists (i.e., not falling asleep/not staying asleep), of which we know (2) how it affects our three measurements (e.g., trouble with falling asleep will be measured as a long time lying awake before falling asleep for the first time during a nightly observation in the laboratory); and we know (3) that it explains the correlation between the three measurements (i.e., the common cause of measures obtained in a sleep laboratory and of ticking the box "long time to fall asleep" in a questionnaire).

In case of mental disorders, on the other hand, a latent variable model is an unlikely candidate for giving a truthful explanation of the associations between distinct symptoms of a disorder. In other words, we do not object to measurement models per se, but to the idea that the association between a mental disorder and its symptoms is one of measurement. First, many supposed latent variables in psychological science—such as depression or neuroticism—do not appear to have a natural referent (for an elaboration on this point, see Borsboom, Cramer, Kievit, Zand Scholten, & Franic, 2009). Second, without a natural referent, we have no idea how the supposed measurements would be affected by the latent variable, and we therefore cannot justify a common cause interpretation, where the disorder explains correlations between its symptoms. Thus, the things that render the correlation between insomnia and three observed variables one of measurement are lacking in the case of, say, depression. Naturally, if one day we should find a natural referent for the hypothetical construct "depression", and we could prove that referent to be the common cause of all depression symptoms, the network model would be disproved. But we doubt that day will ever come.

Refining the network approach: Genetics, brain, and cognition

The network model in Chapter 2 is, naturally, not the end of the story (**Ross**). To the contrary, the network we presented for comorbidity between major depression and generalized anxiety represents a starting point. Refining this model in particular—and the network idea in general—should be the focus of future research in order to adequately

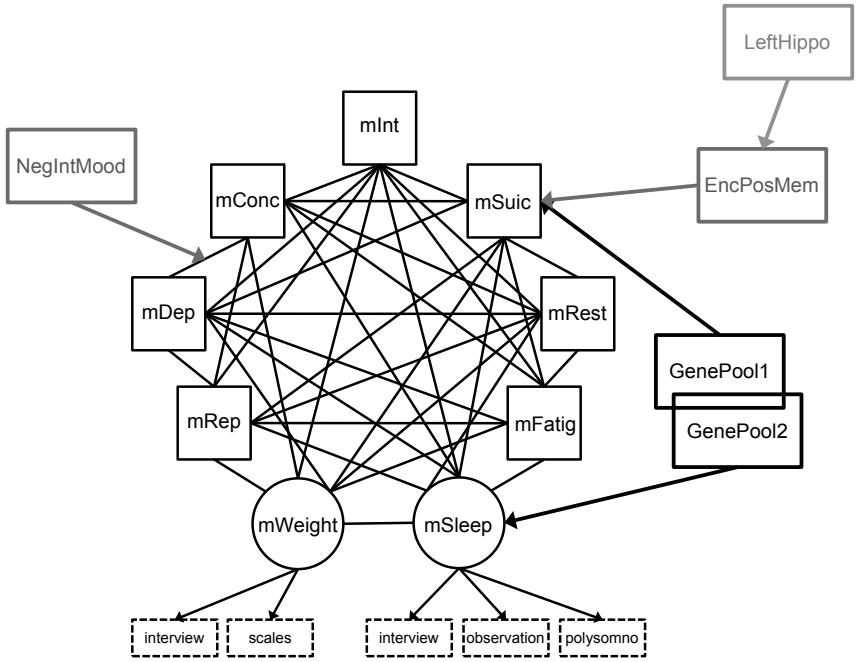

Figure 3.1: A hypothetical network model for major depression. *Circles* represent latent variables and *squares* and *rectangles* represent observed variables. The nine symptoms of major depression are represented as *black squares/circles*. The *dashed rectangles* represent multiple measurements for latent symptoms (i.e., weight and sleep problems in this example; see Section 2). The *black rectangles* represent the genetic level, the *light grey rectangle* the neurophysiological level and the *medium grey rectangles* the cognitive level of the model. *mInt*: loss of interest; *mDep*: depressed mood; *mFatig*: fatigue; *mRest*: restlessness; *mSleep*: sleep disturbances; *mWeight*: weight problems; *mSuic*: (thoughts of) suicide; *mRep*: self-reproach; *mConc*: concentration problems; *polysomno*: polysomnography; *EncPosMem*: problems in encoding/retrieving positive autobiographical memories; *NegIntMood*: negative interpretation of bad mood; and *LeftHippo*: smaller volume of the left hippocampus.

(1) test the validity of the model and (2) generate hypotheses about the etiology of particular mental disorders (**Johnson & Penke**).

Johnson and Penke correctly state that an important goal of the network model is to help unravel the etiology of a wide variety of mental disorders. We acknowledge that a plethora of work has already been done in that regard, but, as we also argued in Chapter 2, that work might be grounded in the wrong psychometric theory of mental disorders. As such, *etiology* is currently interpreted in terms of the development of a single vulnerability (i.e., the common cause) that causes a cluster of symptoms. For example, an evolving lack of serotonin may be hypothesized to cause the symptoms of major depression. However, if a network approach, rather than a latent variable model, correctly describes the system, the conceptualization of etiology and vulnerability radically changes, for we are no longer talking about one, but about a multitude of vulnerabilities at the genetic, neurological, and cognitive levels that may explain the onset of symptoms and the relationships between them (**Fleeson et al.; Hyland; Rubinsten & Henik; Yordanova et al.**). Figure 1 depicts such a hypothetical descriptive network model for the nine symptoms of major depression. The etiology may then be conceptualized in terms of the development of such a network over time; naturally, this process may differ

over individuals.

Many mental disorders have a strong *genetic* component, as evidenced by high heritability estimates, but, despite numerous research efforts, the genetic culprits have not been found (**van der Sluis et al.**). This poses a dilemma. Are the heritability estimates wrong—and is the genetic influence on mental disorders hence highly exaggerated—or is there something wrong with the methods we use to investigate this issue? Van der Sluis et al. suggest the latter and corroborate this by referring to the practice of correlating genes to the entire aggregate of symptoms. If the network model is accurate in describing the origins of mental disorders, this method provides limited prospects for success in gene hunting. Since, in this case, there simply is no common cause, its hypothesized proxy (i.e., a sum score) is an amalgam of distinct factors and will only capture the genetic components that are shared by the aggregated symptoms and relations between them. As we have argued in Chapter 2, it is likely that different genes (or constellations of genes) influence different symptoms (and relations between them). For instance, it is not a wild guess to assume that the symptoms "sleep disturbances" (*mSleep*) and "thoughts of suicide" (*mSuic*) are controlled by a different set of genes (with some overlap; see Figure 1). Multiple genes for each symptom separately does render the entire picture far more complex and we agree with van der Sluis et al. that the network model faces a challenge in that regard. Part of this complexity could possibly be tackled by examining the time series of symptom development and relating the patterns that emerge from such analyses to (constellations) of genes.

While we generally reject the idea of one common cause underlying a constellation of symptoms, we by no means dismiss the potential relevance of pathological mechanisms discovered by the quest of finding such causes. For example, a smaller left hippocampal volume has been consistently found in people with major depression (e.g., see Bremner et al., 2000). Although it appears unlikely that this mechanism causes all depression symptoms, it could be one of the vulnerabilities underlying one or more symptoms; for instance, thoughts of suicide (see Figure 1). Also at the *neurological* level, **Rubinsten and Henik** argue that abnormalities of the intraparietal sulcus (IPS)—commonly associated with numerical cognition—are the common cause of the symptoms of developmental dyscalculia (DD). Although we agree that the evidence points to the relevance of IPS deficiencies, we are not so sure that those deficiencies are the common cause. Since DD involves deficiencies in a variety of complex abilities that require input from memory, attention, and spatial systems, a single underlying vulnerability is highly unlikely (e.g., see Cohen Kadosh & Walsh, 2009; Landerl, Bevan, & Butterworth, 2004). Thus, also in the case of DD, existing neurophysiological findings can be incorporated easily into a network perspective once one is willing to accept the demise of the "common cause" idea.

At the *cognitive* level, it is, for instance, well known that both major depression and generalized anxiety are intimately connected to negative beliefs, as is evidenced by the success of cognitive therapy in reducing depression and preventing relapse (DeRubeis et al., 2005; Kuyken et al., 2008; Papageorgiou & Wells, 2001; Paykel et al., 1999; Wells & Carter, 2001) (see also **Hyland**). We are skeptical about Hyland's view that those beliefs form an interconnected system that completely explains the onset of depression and/or generalized anxiety. Rather, we hypothesize that negative beliefs directly influence (1) symptoms —for example, negative thinking that causes a depressed mood; and (2) relations between symptoms —for example, an overly negative interpretation of one's depressed mood that results in making a suicide plan (see Figure 1). **Stanilou and Markowitsch** report another intriguing possibility: Problems in encoding and retrieving positive autobiographical memories could result in an inability to imagine an optimistic future, which may lead to the onset of the symptom "suicide attempt" (Markowitsch,

Vandekerckhove, Lanfermann, & Russ, 2003; Schacter & Addis, 2009; Sharot, Riccardi, Raio, & Phelps, 2007).

How to investigate the network model? A research agenda

We have provided several arguments for the thesis that a network model paints a more realistic picture of mental disorders than the latent variable model does. Naturally, future research must determine whether the network model is also the better model in reality, and several commentators have put forward some excellent suggestions for a research agenda (e.g., **Davis & Plomin; Fleeson et al.; Tzur-Bitan, Meiran & Shahar**). Given the complexity of the network approach, such an agenda is necessarily comprehensive. As such, when **Krueger et al.** ask, "How would one use the information in Figure 2.4 to explain to a policy maker how we might go about spending public funds wisely in the service of working to ameliorate the burden of depression and anxiety? By funding hundreds of separate projects focused on understanding each line in the figure?" —our short answer is yes. For those skeptical of this answer, we suggest that the same question may be asked about, say, complex systems like the earth's climate. Should we really fund hundreds of projects investigating the diverse factors that influence climate change? The answer to that question is uncontroversially affirmative, and it has not proven difficult to persuade policy makers of this fact. We do not see why the situation would be different for mental disorders. Given this perspective, we think of three lines along which network research should ideally be aligned: (1) validating the network model, (2) elucidating the vulnerabilities underlying (relations between) symptoms (see **Fleeson et al.**) and (3) tracking the developmental trajectories of symptom constellations.

Validating the network model

Relations between symptoms represent an ideal opportunity to test the network model against the latent variable model: If no latent variable exists, one should find that experimentally manipulating one symptom results in change in another symptom. Some work has already been done in that regard; for example, unsurprisingly, one look at the literature reveals a direct effect of sleep deprivation on fatigue (e.g., see Durmer & Dinges, 2005). Other symptom relations, such as the one between loss of interest and worrying about multiple events in Figure 2.4 of Chapter 2, appear less obvious and need experimental verification in the future. In a more direct manner, the network model could be confirmed by the genetic association studies (GAS) on the individual symptoms, as proposed by **Van der Sluis et al.**; it would be especially interesting to execute such analyses on patterns found in time series that describe symptom dynamics. If the network model is true, this type of GAS should reveal constellations of genes that better account for the high heritability of mental disorders than GAS on a sum score. In the same vein, Davis and Plomin suggest multidimensional scaling as a method to reveal the genetic closeness of multiple symptoms. If such endeavors would point to the presence of direct relations between symptoms, the latent variable model could be put to rest in psychopathology.

Elucidating vulnerabilities

Fortunately, there may be no need for funding "hundreds of projects", as **Krueger et al.** fear, since many of such projects, aimed at understanding the inner workings of a variety

of symptoms, have already been carried out; most symptoms in Figure 2.4 in Chapter 2 are associated with large scientific literatures (e.g., fatigue, anxiety). With regard to vulnerabilities underlying the relations between symptoms, not all edges are an a priori mystery to us; for example, the mechanisms that are involved in the influence of sleep deprivation on fatigue are quite well-known (e.g., see Durmer & Dinges, 2005).

With regard to symptom relations whose underlying mechanisms are less well-known, insights from treatment rationales should further our understanding. For instance, mindfulness based cognitive therapy offers a specific hypothesis with regard to the relation of depressed mood with the other symptoms of depression: Depressed mood triggers ruminative thinking, which—if not hindered by a successful intervention—could lead to other depression symptoms (e.g., see Ma & Teasdale, 2004; Nolen-Hoeksema, 2000; Segal, Williams, & Teasdale, 2002). Another example comes from the panic disorder literature in which renewed interpretation of bodily signals is used to break the link between having a panic attack and worrying about its consequences ("I will have a heart attack"; e.g., see D. M. Clark et al., 1994). On a related note, several successful interventions are not primarily aimed at reducing or eliminating symptoms or the relations between them but, rather, at reinforcing so-called protective factors. For example, the relative success of the methadone program is attributable to reinforcing coping skills and finding work and housing (i.e., protective factors) while stabilizing the addiction with the methadone. Once a stable situation is created, addicts enter a total abstinence program (e.g., Gossop, Stewart, Browne, & Marsden, 2002; van den Brink, Hendriks, & van Ree, 1999). Such treatment programs could provide some valuable insights into the mechanisms by which one progresses from a disordered to a healthy state.

Tracking developmental trajectories

Much of the current literature reports research that involves interindividual research, often carried out cross-sectionally. Although such research can provide important insights, **Wass and Karmiloff-Smith** correctly suggest that it results in a snapshot of reality: an interindividual picture of mental disorders, frozen at a particular time frame. In reality, it is likely that, for instance, edge strengths differ across individuals, as well as across time. If so, another line of research is required to generate answers to two pivotal questions: (1) How do mental disorders develop, and (2) how does that development differ across individuals (**Fleeson et al.; Rothenberger et al.**). Such differences should be detectable through the intra-individual analysis of time series, as noted by various commentators (e.g., Fleeson et al.; **Tzur-Bitan et al.; van Geert & Steenbeek**). In earlier times, it was quite difficult to obtain data suitable for such analyses. Fortunately, we now live in a time in which intensive time series data can be gathered relatively easily (e.g., by letting patients report the status of symptoms through handheld devices, etc.). We think that, within a few years, it will become possible to analyze symptom development in real time, and to update network structures and parameters as the data come in. And when that time comes, we are confident that thorough investigation of the network approach will result in a better understanding of symptoms, their relationships, and their course in individuals over time.

Extending the network approach to other disorders

In Chapter 2, we introduced the network approach for two disorders that are prime examples of Axis I disorders in the DSM-IV (APA, 1994). Any theory that presents itself as *the* potential substantive theory of mental disorders must be able to explain more than comorbidity between major depression and generalized anxiety disorder (**Johnson**

& Penke). As a first step, we deem it necessary to evaluate to what extent the network approach fits a variety of other mental disorders (also see **Cervone**).

With regard to other Axis I disorders, some commentators have presented specific examples of (clusters of) disorders for which common causes are supposedly identified, thereby rendering the network approach invalid in those cases (e.g., **Ross; Rubinsten & Henik**). For example, Ross argues that addictions share a common cause: namely, hyperactivation of the dopamine reward circuit combined with weakened frontal and prefrontal serotonin and gamma-aminobutyric acid (GABA) circuits. We share Ross's view on the importance of these brain pathologies in addiction; however, we do not agree that such pathologies automatically qualify as the common cause of addictions. The most commonly reported consequences of the dysfunctional dopamine, serotonin, and GABA circuits are (1) the strong desire to consume salient targets, coupled with (2) difficulty resisting that desire. In other words, the brain pathologies that Ross mentions result in the *core characteristics* of an addiction. However, does this make those brain pathologies the *common cause* of addiction? To qualify as such, those pathologies should *cause* the other symptoms of addiction. This is unlikely.

If we take a look at the DSM-IV criteria for substance abuse, for instance, we notice (1) the apparent inability of dysfunctional neurotransmitter circuits to explain "recurrent substance use *resulting in a failure to fulfill major role obligations*"; and (2) the undeniable possibility of direct relations between the symptoms of addiction: "Recurrent substance use in situations in which it is physically hazardous" (e.g., drunk driving) can cause "recurrent substance-related legal problems" (e.g., getting arrested for drunk driving). As such, we think addiction can potentially be envisioned as a causal chain of symptoms in which one symptom—the desire to consume a substance and the inability to withstand this—may be triggered by dysfunctional dopamine, serotonin, and GABA circuits; thus, no common cause, but one pathological mechanism—in combination with other etiological factors—potentially results in a cascade of events in a network of addiction symptoms (i.e., the "fan-out" principle that **Wass & Karmiloff-Smith** mention). Such a chain of symptoms is also likely in panic disorder and other—very heterogeneous—Axis I disorders such as schizophrenia and attention-deficit hyperactivity disorder (ADHD). Hence, in these cases the network approach cannot be ruled out a priori (e.g., Borsboom, 2008b).

Considering the extension of the network approach to Axis II disorders, **Bornstein** sees some roadblocks that need to be overcome in the case of personality disorders (PDs). First, patients with PDs tend to experience their symptoms as congruent with themselves. As a result, these patients have limited insight into their own condition. Bornstein rightly sees two resulting consequences: (1) Self-report measurements alone will not be adequate in assessing people with suspected personality pathology, and (2) the symptoms that patients cannot reflect on themselves are in a sense "latent". However, we do not think these consequences pose serious problems for the network approach since—as we outlined in Section 2 of this chapter—it can easily deal with latent variables that have an established measurement relationship with a set of indicators, including tests that do not rely on self-assessment. Second, the revision of PD symptoms is founded on a desire to both increase diagnostic accuracy and reduce comorbidity. According to Bornstein (2003), this practice has resulted in simply removing symptoms from the diagnostic checklist, and, as Bornstein rightly claims, this poses a potential problem for the network approach; however, not in terms of its potential as substantive theory of mental disorders, but in terms of its practical applicability to PDs with potentially incomplete symptom inventories. So, in the case of Axis II PDs, we see no immediate problems that the network approach cannot surmount.

What is a mental disorder?

In Chapter 2, we argued that boundaries between mental disorders are necessarily fuzzy. In contrast, **Haslam** argues that boundaries between categories of the same disorder (e.g., "disordered" versus "not disordered") are not fuzzy at all. To address this apparent dilemma properly, we dissect a disorder network in two components: (1) its structure and (2) its state. The *structure* of a disorder network refers to the strength of the relations between symptoms. As we show in Figure 1, these relations are controlled by a host of vulnerabilities (e.g., negative interpretation of one's mood resulting in a relatively strong relation between depressed mood and thoughts of suicide). Since those vulnerabilities probably differ across individuals, it is safe to assume that the resulting basic network structure is individually tailored as well. Now, pertaining to comorbidity, it is likely that, in some cases, individual network structures do not obey the DSM boundaries between disorders (nor any other fixed boundaries). It is likely as well that certain vulnerabilities influence relations between symptoms of different disorders: for instance, ruminative thinking may strengthen the relation between "depressed mood" and "chronic anxiety". As such, the boundary between major depression and generalized anxiety for someone with a ruminative thinking style probably (1) does not equal the DSM-defined boundary (because of a strong relation between "depressed mood" and "chronic anxiety") and (2) lies somewhere else than the boundary of someone without that thinking style. Thus, at the individual level, the line can be drawn practically anywhere and therefore we defend the notion of fuzzy boundaries in these cases. In other cases, a sharp boundary might be more feasible; for instance, because relations between symptoms of these disorders are virtually nonexistent or negative. For example, large individual differences in the boundary between social anxiety and psychopathy are not very likely given the opposite nature of the symptoms of those disorders (e.g., "excessive self-consciousness and anxiety in everyday social situations" versus "grandiose sense of self-worth"; Hare, 2003).

The *state* of a disorder network depends on how much symptoms are "on". When adhering to a categorical perspective, disorder networks can be in two or more stable states. For example, with two stable states, one commonly distinguishes between a *healthy state*, in which few symptoms are "on", and a *disordered state*, in which several symptoms are "on". In these cases, a sharp boundary is needed to distinguish few from several. Now, we agree with **Haslam** that such sharp boundaries are theoretically possible and that evidence for two latent classes corroborates that hypothesis (provided that the analysis was conducted on a large and representative sample). However, as we already argued for the structure of a network, it is unlikely that boundaries between states are invariant over persons; for, in subjective terms, some people feel depressed because they have sleep and concentration problems for two weeks, whereas others succumb to a full-blown depression only after a prolonged period of experiencing a multitude of symptoms. Therefore, in these cases, a more dimensional perspective might be in order; that is, no sharp boundaries between categories, but, instead, a continuum of network activation. Here, we think that symptom severity might be an excellent candidate for representing the degree of network activation (**Markus**): the more severe someone's symptoms, the more that person is located at the "disordered" end of the continuum.

In theory, any network with connected nodes (i.e., structure) that can be in different states could be taken to qualify as a mental disorder. As such, liberalism could be viewed as a mental disorder (**Zachar**): a set of connected political beliefs (e.g., if you believe in freedom of religion for everyone, then it is more likely that you are tolerant of minorities) that we call "liberalism" when a sufficient number of nodes are activated. In practice, though, we—and probably the majority of humankind with us—do not consider liberalism to be a mental disorder. Why? The DSM provides a sensible answer: The symptoms of any candidate mental disorder should cause "clinically significant distress

or impairment in social, occupational, or other important areas of functioning" in the person who is experiencing these symptoms (APA, 1994). Although liberalism apparently causes distress in some *other* people (Savage, 2005), it clearly does not satisfy the DSM's prerequisite. Thus, providing a sensible boundary between disorders and non-disorders, we would welcome this prerequisite as an extra node in the symptom space.

About 40% of people with major depression experience a new depressive episode after treatment (e.g., Paykel, 2008). Any substantive theory of mental disorders must be able to explain such *recurrence*, a phenomenon that is very common in a host of mental disorders. In our opinion, the network approach is up to that task. Take, for instance, an alcoholic who, because of treatment, manages to stay sober, as a result of which the other symptoms of his/her substance abuse also subside. Also suppose that this person's network has strong connections between symptoms; that is, if one symptom turns on, it is likely that the other symptoms will turn on as well. As such, we have a situation in which the substance abuse network is in a more or less *healthy* state (i.e., no symptoms are "on") while the structure of the network is *risky* (and thus unhealthy). Now, this is exactly what makes a disorder likely to recur: If, for whatever reason, this person decides to drink one beer, it will likely result in a cascade of symptoms being turned on, and eventually the network will return to a disordered state. In other words, recurrence is most likely when the healthy state of a disorder network is unstable because of the strong connections between its symptoms. We think this is precisely what clinicians mean when they talk about silent disorders, and therefore we do not agree with **Hood and Lovett** that the network approach cannot accommodate such notions. On a final note, in the case of major depression, it is established that one of the most reliable predictors of recurrence is the presence of residual symptoms (e.g., Kennedy & Paykel, 2004). But we also know that not every patient with residual symptoms experiences a subsequent recurrence. If we are right in suggesting that recurrence is most likely when the structure of the network is strong, residual symptoms in depression patients offer a way to prove this hypothesis: Of patients with residual symptoms, only those with strong connections between symptoms should eventually experience a new episode of major depression.

Networks versus common causes: Methodological issues

Several commentators raise methodological issues regarding the network approach as opposed to latent variable models. In the following, we discuss criticisms according to the methodological topics mentioned by the commentators.

Local independence

Many commentators question our criticism of the local independence assumption. In their opinion, a unidimensional model with local independence is unnecessarily strict (e.g., **Humphry & McGrane; Markus; Molenaar**). It is true that violations of local independence can be represented in a latent variable model, for instance, by allowing correlated residuals or direct relations between indicator variables. However, these modeling possibilities should not be given too much conceptual weight. Being more than a convenient restriction, local independence has the status of an axiom in measurement models used in psychometrics (e.g., Ellis & Junker, 1997; Holland & Rosenbaum, 1986; Junker & Sijtsma, 2001). This makes sense because psychometric models aim to give conditions under which composite scores (e.g., summed item scores) can be treated as measures of a latent variable. A prerequisite for this is that the item scores measure

the same latent variable, which plausibly requires that the latent variables functions as a common cause; and the classical way of testing this is by testing whether the latent variable screens off the associations between the item scores. This is precisely what local independence requires. Thus, although it is statistically possible to allow for direct relations between indicator variables in a model, this should be considered a deviation from a psychometric norm (which in itself is reasonable in setting up a measurement model). As such, a unidimensional model with local independence is anything but a "straw man" (**Danks et al.**).

Model equivalence

Several commentators raise the possibility that we may have overstated the difference between networks and latent variable models. **Danks et al.** note that cyclic graphs and latent variable models are closely related; **Molenaar** points to the fact that longitudinal factor models are equivalent to specific types of directed network models; and **Humphry and McGrane** indicate that latent variable models concern individual differences and, as such, may allow for individual level causal relations without violating the individual differences model.

It is true that latent variable models and network models are statistically indistinguishable in certain situations. A prominent example of such an exact indistinguishability is the mutualism model of intelligence proposed by Van der Maas et al. (2006), which is a network model that can produce data that are exactly equivalent to a single factor model. Similar relations are likely to exist for item response theory (IRT) models; **Molenaar**, in earlier work (see Molenaar, 2003, p. 82) has noted the close relation between Markov field models, such as the Ising model, and IRT models like those of Rasch (1960) and Birnbaum (1968). Indeed, one supposes that model equivalence may obtain as well in those cases.

Does this render the network model and the latent variable model equivalent in general? No, because the inability to distinguish between different possible generating models in a given data-set does not imply that the models are equivalent with respect to all possible data-sets or under all possible interventions. Thus, the advice in a model equivalence situation is to get better data, such as intensive time series (see Section 4).

Parsimony

Krueger et al. defend the latent variable model by emphasizing its superior parsimony relative to the network approach. First, latent variable models are not *inherently* more parsimonious than network models because the number of parameters of the latter can be made arbitrarily small. For instance, suppose that one has k observed dichotomous symptoms. If one assumes a completely connected network consisting of bidirectional relations of equal size, where these relations are functionally the same for any two nodes (e.g., logistic relations with equal intercepts and slopes, as in a Boltzmann machine; see Ackley, Hinton, & Sejnowski, 1985), then, statistically speaking, one has an extremely parsimonious model even though it may consist of many—namely, $(k(k - 1)/2)$—connections between variables.

Second, it should be recognized that even though parsimony is a useful criterion in choosing between statistical models, it will lead to truth only if reality itself is simple; it this is not the case, then we may deceive ourselves by overemphasizing parsimony. As Tryon (1935, p. 428) remarked, "The 'law' of parsimony is not a natural law, but a rule agreed upon men to simplify their thinking". While simplifying our thinking is clearly useful in scientific investigation, complex realities will ultimately require complex models. In the case of mental disorders, we doubt that reality is simple given the likelihood of

variation in network structure over individuals and time. As such, an extremely restricted model such as Boltzmann machine—although favorable in terms of its parsimony—might not be particularly viable. Therefore, we think that the sword of parsimony should be wielded with caution, for we may accidentally kill promising candidate models through its use.

Extensions of the network approach

Danks et al. provide one of the most critical analyses of our approach. First, they raise a number of questions concerning terminology and procedure. For instance, they criticize our use of the term *centrality* because "[centrality] is neither a causal nor a statistical notion." This is obviously correct; it is a notion that comes from network analysis and has proved to be useful in many contexts (e.g., see Boccaletti et al., 2006). Danks et al. also question our statement that observables in a standard psychometric latent variable model are exchangeable. In a measurement model, observables do not differ with respect to the property they measure; they are thus exchangeable in this sense. And it is this exchangeability that—among other things—renders the standard measurement model inappropriate in the context of psychopathology, for how could "weight loss" measure the same property as "suicide plans"[1]? Finally, Danks et al. indicate that the data we analyzed involved a great deal of missingness. We agree but refer to Footnote 6 in Chapter 2, where we highlight an appropriate estimation approach we used to deal with the data, which is missing at random because of the skip structure of the interview schedule used in the National Comorbidity Survey Replication (NCS-R).

Second, **Danks et al.** state that we "do not engage what is known" about the investigation of causal relations, instead settling for an unsatisfactory and unrestrictive visualization method. They propose that causal inference algorithms should be used instead and report the outcome of an algorithmic search procedure. Perhaps ironically, the use of such procedures formed the starting point of our research. However, the search procedures as implemented in the program TETRAD (Scheines, Spirtes, Glymour, Meek, & Richardson, 1996) returned causal structures that we felt were extremely hard to make sense of. This is also the case for the model suggested by Danks et al., in which, for instance, the core symptoms of depression and generalized anxiety (i.e., depressed mood and chronic anxiety) are completely disconnected from the model. Our diagnosis of this situation is that two assumptions of the search algorithms in existence are not satisfied in the data at hand: (1) Individuals have the exact same causal structure and (2) resulting graphs are acyclic. In contrast, we think that the network structure of mental disorders (1) varies over individuals and (2) likely contains feedback loops. Therefore, we judge the implementation of causal search algorithms to the preliminary; it would be more sensible to gather time series data on symptom dynamics and to fit models on an intra-individual basis. However, what we can do unproblematically, absent such intensive time series data, is to provide a starting point for further investigations and hypothesis formulation,

[1] The exchangeability of items with respect to the property they measure is clear from the fact that one can parameterize, for instance, standard IRT models such as the one- and two-parameter logistic models by identifying the latent variable with the expectation of any one of the item responses (Gunter Maris, personal communication). A similar situation holds for the (essentially) tau-equivalent model of classical test theory (Lord & Novick, 1968), in which the expectations of observed variables are simple transformations of one another, and for the congeneric model of factor analysis, in which the observed variables are linear transformations of one another (Jöreskog, 1971)). Intuitively, this means that if one has a single perfect thermometer, adding information from other, noisy thermometers is useless (note that this makes sense in a measurement situation). In contrast, if one knew the expectation of the item "how much weight have you lost?" one would presumably still want to know whether the person had suicide plans.

based on the visualization of statistical associations that exist in the data, and this is what we aimed to do. This does not commit us to any particular type of modeling, while it serves the purpose of introducing and explaining the network approach extremely well. In conditions that justify their use, however, we acknowledge that causal modeling and search algorithms may be very useful.

Chapter 4

The relation between stressful life events and depressive symptoms

Abstract

Previous research has shown that stressful life events (SLEs) influence the pattern of individual depressive symptoms. However, we do not know how these differences arise. Two theories about the nature of psychiatric disorders have different predictions about the source of these differences: 1) SLEs influence depressive symptoms and correlations between them indirectly, via an underlying acute liability to develop a depressive episode (DE; common cause hypothesis); and 2) SLEs influence depressive symptoms and correlations between them directly (network hypothesis). The present study investigates the predictions of these two theories. We divided a population-based sample of 2096 Caucasian twins (49.9% female) who reported at least two aggregated depressive symptoms in the last year into four groups, based on the SLE they reported causing their symptoms. For these groups, we calculated tetrachoric correlations between the 14 disaggregated depressive symptoms and, subsequently, tested whether the resulting correlation patterns were significantly different and if those differences could be explained by underlying differences in a single acute liability to develop a DE. The four SLE groups had markedly different correlation patterns between the depressive symptoms. These differences were significant and could not be explained by underlying differences in the acute liability to develop a DE. Our results are not compatible with the common cause perspective but are consistent with the predictions of the network hypothesis. We elaborate on the implications of a conceptual shift to the network perspective for our diagnostic and philosophical approach to the concept of what constitutes a psychiatric disorder.

Adapted from: Cramer, A. O. J., Borsboom, D., Aggen, S. H., & Kendler, K. S. (2012). The pathoplasticity of dysphoric episodes: differential impact of stressful life events on the pattern of depressive symptom inter-correlations. *Psychological Medicine, 42*, 957-965.

Depressive symptom profiles of people were long thought to be stable. Until well in the 1990s mainstream theories (e.g., diathesis-stress model) predicted that both between and within individuals, the pattern of symptoms displayed across multiple episodes is roughly the same. As such, it was thought that external factors, like stressful life events (SLEs), that were known to be associated with an increased risk for the onset of an episode of major depression (MD; e.g., Kendler, Karkowski, & Prescott, 1999; Rijsdijk et al., 2001; Leskelä et al., 2004; Olsen, Mortensen, & Bech, 2004; Jacobs et al., 2006; Middeldorp, Cath, Beem, Willemsen, & Boomsma, 2008; Munafò, Durrant, Lewis, & Flint, 2009), were not capable of influencing the occurrence of individual symptoms. This assertion changed when research showed that depressive symptom profiles across multiple episodes of MD within the same individual were moderately stable at best (e.g., Coryell et al., 1994; Oquendo et al., 2004): could this partly be due to the direct influence of SLEs on individual symptoms? Yes, depressive symptoms were shown to vary as a function of the particular class of SLEs that preceded the onset of these symptoms (Keller & Nesse, 2005, 2006; Keller, Neale, & Kendler, 2007; Slavich, Thornton, Torres, Monroe, & Gotlib, 2009: for example, romantic breakups were associated with high levels of guilt, while stress was associated with fatigue and hypersomnia (Keller et al., 2007). Thus, depressive symptom profiles are more pathoplastic than was once thought in that environmental precipitants like SLEs can give 'content, coloring and contour' to the individual expression of such profiles (K. Birnbaum, 1923). Instead of individual symptoms, this chapter presents a novel approach in which the impact of SLEs on the overall pattern of correlations between these symptoms is investigated.

Why is studying the impact of SLEs on correlations between depressive symptoms important? An answer brings us back to the pioneering work of, for example, Kraepelin (1923) who tried to distinguish between psychiatric disorders based on the observation that some symptoms were more often seen together in patients than others. For example, depressed mood and feelings of worthlessness were displayed in patients more frequently than depressed mood and disorganized thinking. Many similar observations later culminated in the definition of distinct psychiatric disorders, designating depressed mood and feelings of worthlessness as symptoms of MD and disorganized thinking and thought insertion as symptoms of schizophrenia. Put in statistical system, the setup of the current classification system is based on the fact that some symptoms are more strongly correlated with each other (e.g., depressive symptoms with one another) than with other symptoms (e.g., depressive symptoms with symptoms of schizophrenia).

The critical question is why symptoms of psychiatric disorders are strongly intercorrelated. The leading hypothesis—one that is more often assumed than examined critically—postulates a common cause framework (Pearl, 2000; Bollen, 2002; Borsboom et al., 2003; Reise & Waller, 2009). Panel A of Figure 4.1 displays how correlations between symptoms of a depressive episode (DE), that is, an episode of MD, would be understood from this perspective: the DE is the common cause of its symptoms. That is, depressive symptoms are correlated because they are caused by the same underlying (acute) liability to develop a DE. Importantly, this perspective claims that correlations between symptoms are not indicative of a real relationship between them: insomnia (or hypersomnia), for example, is not directly related to fatigue; both symptoms are only correlated because they are both caused by the same underlying depressive liability. Recently, a novel alternative has been articulated (Cramer, Waldorp, van der Maas, & Borsboom, 2010) in which there is no common cause (see panel B of Figure 4.1). Instead, correlations between symptoms (lines between the symptoms in the figure) represent real relationships (possibly causal in nature) and, as such, the connected symptoms form a network. That is, this alternative model postulates that a DE (and its more severe counterpart, MD) is a network of symptoms that stand in direct (causal) relations

toward one another. The most compelling argument for a network account of psychiatric disorders is commonsensical. It seems unrealistic to assume that insomnia and fatigue are only correlated because both are a result of an underlying liability to develop a DE. Surely, all of us have experienced that having trouble sleeping can directly lead to tiredness the next day.

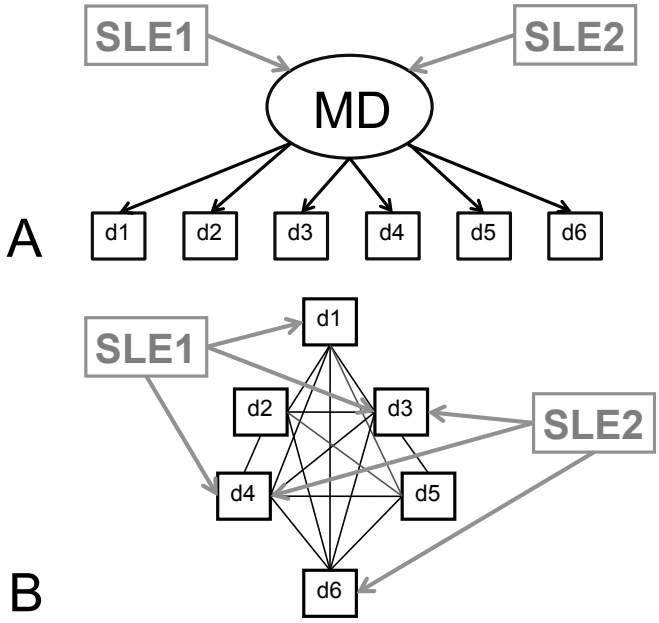

Figure 4.1: Major depression (MD) according to a common cause (A) and a network (B) perspective. A: the common cause MD causes its six symptoms (d1-d6). Stressful life events (SLE1 and SLE2) influence the symptoms of MD only indirectly, via the common cause MD. B: MD is a network in which symptoms d1-d6 are directly connected with one another. SLE1 and SLE2 influence the symptoms of a depressive episode directly.

These two ways of conceptualizing psychiatric disorders assume a different relationship between a DE and SLEs. According to a common cause perspective (see panel A of Figure 4.1), SLEs influence the symptoms only indirectly, via their impact on the common cause: if, after SLE1, d3 and d4 are more strongly correlated than after SLE2, this is because the acute liability to develop a DE (i.e., common cause) is increased after SLE1 compared with SLE2. According to a network perspective (see panel B of Figure 4.1), SLEs can influence symptoms directly: if, after SLE1, d3 and d4 are more strongly correlated than after SLE2, this results from a real increase in the strength of the correlation between d3 and d4 after SLE1 compared with SLE2 due to the direct impact of both SLE1 and SLE2 on depressive symptoms. Hence, if SLEs make an impact on the pattern of correlations between depressive symptoms differently, the common cause perspective predicts that those differences are due to underlying differences in acute liability to develop a DE while the network perspective predicts a direct influence of SLEs on depressive symptoms and correlations between them. This chapter investigates the prediction of the common cause perspective by a comparison of the impact of four SLEs

on disaggregated depressive symptoms[1] in twins from a general population sample with a DE in the last year.

Method

Participants

The data for this chapter consisted of a subsample of 2096 participants (49.9% female) from the larger Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPUD), a population-based longitudinal study of Caucasian twins from the Mid-Atlantic Twin Registry (for details, see Prescott, Aggen, & Kendler, 2000; Kendler, 2006). The present chapter is based on member of female-female, male-male and male-female twin pairs who, at the first interview, reported (1) having experienced a DE (see Measures section) and (2) that their depressive symptoms were precipitated by one of four SLEs.

Measures

The first VATSPUD interview assessed the presence/absence of the 14 disaggregated symptoms of MD (representing the nine aggregated symptoms of criterion A for MD in DSM-III-R), lasting at least 5 days during the previous year. Whenever a symptom was present, interviewers probed to ensure that its occurrence was not due to medication or physical illness. Participants were then asked which symptoms co-occurred, and the interviewer aggregated these symptoms into syndromes. Following earlier work (Keller et al., 2007) we define a DE as any syndrome in which two or more of the nine aggregated depressive symptoms co-occurred.

For each DE, participants were asked whether something had happened to make them feel that way or the symptoms just came out of the blue: this methodology is to a substantial extent based on the Life Events and Difficulties measures (LEDS; G. W. Brown, Bifulco, & Harris, 1987) with the main difference that the LEDS concepts were adapted to be rated by the interviewer. If participants could think of a reason, they were asked to describe it. Timing of the event was recorded as described by the respondent and if unsure, the interviewer helped them with other key events in the last year. Interviewers subsequently encoded the responses to specific causal codes. If participants indicated multiple causes, they were asked to order them by causal importance. Following earlier work (Keller et al., 2007), the primary codes (i.e., codes that participants had indicated were of highest causal importance) were collapsed into nine SLE groups of which the four most prevalent were used in the analyses: (1) Stress: stress due to work, finances, legal problems, etc.; (2) Romantic loss (henceforth abbreviated as *RomLoss*): ending of a romantic relationship, including divorce; (3) Health: one's own health problems; and (4) Conflict: interpersonal conflict between self and another. Inter-rater reliability for determining the occurrence and dating of the SLEs was found to be in the good to excellent range (see Kendler et al., 1995).

[1] We opted for the disaggregated symptoms instead of the more commonly used aggregated symptoms because correlations between aggregated symptoms are more difficult to interpret. For example, what does a high correlation between sleep disturbances and depressed mood mean: an association between insomnia and depressed mood or between hypersomnia and depressed mood?

Statistical analysis

We computed tetrachoric correlations between the symptoms for each SLE and presented the resulting networks graphically. We conducted three main analyses. Descriptive in nature, the first analysis investigated differences between the four SLE groups in the graphical representation of their symptom networks: for example, is the correlation between depressed mood and thoughts of death stronger in one SLE group compared with the other SLE groups? Also, we analyzed differences between SLE groups by computing each symptom's centrality in their respective networks (i.e., the sum of all tetrachoric correlations between that symptom and all others in a network; Boccaletti et al., 2006): the higher the centrality of a symptom, the more strongly that symptom is connected with other symptoms in the network.

In the second analysis, we tested whether the observed patterns of correlations among the symptoms in the four SLE groups were significantly different from one another. To this end, we assessed whether constraining correlations to be equal across SLE groups (i.e., homogeneity) would result in a poorer relative fit compared with allowing the free estimation of correlations in each SLE group (i.e., heterogeneity). If so, heterogeneity would thus be preferred over homogeneity, and this implies that the differences in the correlation networks between SLE groups are significant.

In the third analysis, we sought to evaluate whether the differences in the correlation networks between the SLE groups could be due to underlying differences in acute liability to develop a DE. To this end, we compared two versions of the model as it is depicted in panel A of Figure 4.1. The first model (model I) assumes that differences in the networks cannot be explained by underlying differences in acute liability to develop a DE (i.e., the impact of different SLEs on the DE circle in the figure is the same). Instead, differences in the networks are explained by differences in the strength of the associations between a DE and its symptoms (i.e., arrows between DE and d1-d6 in the figure). The second model (model II) assumes that differences in the networks can be explained by underlying differences in acute liability to develop a DE: SLEs influence the DE circle in the figure differently while the associations between a DE and its symptoms (i.e., arrows between DE and d1-d6 in the figure) are the same for individual SLEs. We compared the fit of both models[2]: if model II does not fit worse than model I, this would be consistent with a common cause perspective on DEs.

We estimated models in Mplus 4.2 (Muthén & Muthén, 2007) with the weighted least squares mean and variance adjusted estimator and a Delta parameterization. The fit of the models was assessed with (1) the χ^2 statistic (with $0 \leq \chi^2 \leq 2$ degrees of freedom (df) indicating good fit and $2 \leq \chi^2 \leq 3$ df indicating acceptable fit); (2) root mean square error of approximation (RMSEA; with RMSEA ≤ 0.06 indicating good fit); and (3) the comparative fit index (CFI; with CFI ≥ 0.95 indicating good fit; Hu & Bentler, 1999).

Results

Sample characteristics

Descriptive characteristics of the groups of participants exposed to the four different SLEs are provided in Table 1. The average symptom sum score (i.e., the total number

[2] In more technical psychometric terms, model I is a baseline model in which loadings and thresholds are freely estimated in all groups (factor means fixed to 0, and factor variances and scaling factors fixed to 1 in all groups). Model II is a weak factorial invariance model in which loadings are constrained to be equal across groups (factor means and variances freely estimated in all but the first group, scaling factors fixed to 1 in all groups, thresholds freely estimated in all groups).

of endorsed symptoms) differed across the SLE groups (e.g., the average sum score was lowest in the Stress group and highest in the Conflict group) and these differences were significant (non-parametric Kruskal-Wallis test: $\chi^2 = 99.03$, df = 3, $p < 0.001$).

Table 4.1: Descriptive characteristics of the four SLE groups

	Participants (n)	Proportion female (%)	M	SD	Mean rank
Stress	710	42.54	3.80	2.03	880.65
RomLoss	528	50.38	4.88	2.43	1167.91
Health	371	55.53	4.33	2.19	1035.78
Conflict	487	56.06	4.89	2.39	1173.45

SLE, stressful life event; proportion female, percentage of females in (sub)sample; M, average symptom sum score; SD, standard deviation of the average symptom sum score; Stress, stress due to work, finances, legal problems, etc.; RomLoss, ending of a romantic relationship, including divorce; Health, one's own health problems; Conflict, interpersonal conflict between self and another.

Graphical representation of the correlation networks

Figure 4.2 presents the correlation networks for the SLE groups (figure made with the R-package *qgraph*; Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012). In general, the correlations between depressive symptoms are stronger after RomLoss and Conflict than after Stress or Health (where thickness of the connections between symptoms reflects the magnitude of the correlation). This general difference is, however, modest: average correlations between depressive symptoms in the four groups are 0.23, 0.21, 0.19 and 0.17 for the Stress, RomLoss, Health and Conflict groups, respectively. Differences are more substantial when differences in individual correlations between the SLE groups are examined. For example, the correlation between depressed mood (*depr*) and thoughts of death (*deat*) is much stronger in the Health and Conflict groups than in the Stress and RomLoss groups. The connection between feelings of worthlessness (*wort*) and thoughts of death (*deat*) is stronger in the Stress and RomLoss groups than in the Health and Conflict groups. There are also some similarities: in all four groups, weight loss (*wlos*) and weight gain (*wgai*) are strongly connected to decreased appetite (*dapp*) and increased appetite (*iapp*), respectively.

Centrality of MD symptoms in the networks

Figure 4.3 presents the centrality of each symptom in the four SLE groups. Some differences between the groups are worth noting: decreased appetite (*dapp*) is highly central in the Conflict group (is a distinct peak in the graph) while relatively peripheral in the RomLoss group (no peak in the graph). Loss of interest (*inte*) is very central in the RomLoss group while relatively peripheral in the Health group. Finally, fatigue (*fati*) is relatively central in the RomLoss group while relatively peripheral in the Stress and Health groups. In general, differences between SLE groups are more striking than their similarities but one similarity does stand out: feelings of worthlessness (*wort*) and thoughts of death (*deat*) rank among the most central symptoms in every SLE group.

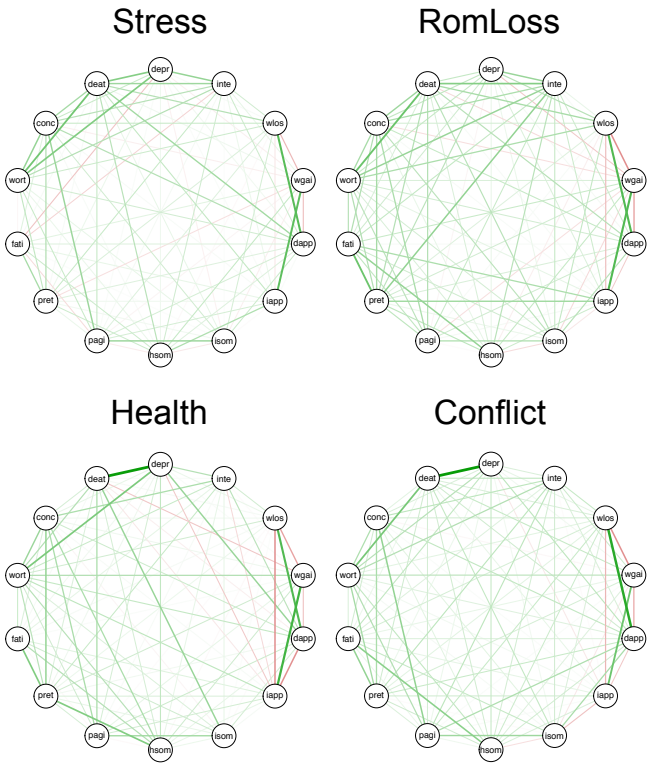

Figure 4.2: Correlation networks between the symptoms of a depressive episode for the four stressful life event groups. The top left depicts the network after stress (Stress); the top right after a romantic loss (RomLoss); the bottom left after health problems (Health); and the bottom right after an interpersonal conflict (Conflict). Each symptom is represented as a node in the networks and a connection between two symptoms represents the tetrachoric correlation between them. The connection is green when the correlation is positive and red when the correlation is negative. *depr*: depressed mood; *inte*: loss of interest; *wlos*: weight loss; *wgai*: weight gain; *dapp*: decreased appetite; *iapp*: increased appetite; *isom*: insomnia; *hsom*: hypersomnia; *pagi*: psychomotor agitation; *pret*: psychomotor retardation; *fati*: fatigue; *wort*: feelings of worthlessness; *conc*: concentration problems; *deat*: thoughts of death.

Homogeneity versus heterogeneity of the correlation networks

The solution in which correlations were estimated separately for each SLE group (i.e., heterogeneity) fitted much better than a solution in which correlations were constrained to be equal across SLE groups (i.e., homogeneity). This result was found in two separate analyses: (1) thresholds were estimated separately in each SLE group in both solutions; and (2) thresholds were constrained to be equal across SLE groups in both solutions. In both analyses, a highly significant χ^2 difference test ($\chi^2 = 384.41$, df $= 273$, $p < 0.001$) indicated heterogeneity: the patterning of tetrachoric correlations between depressive symptoms of the SLE groups is significantly different from one another.

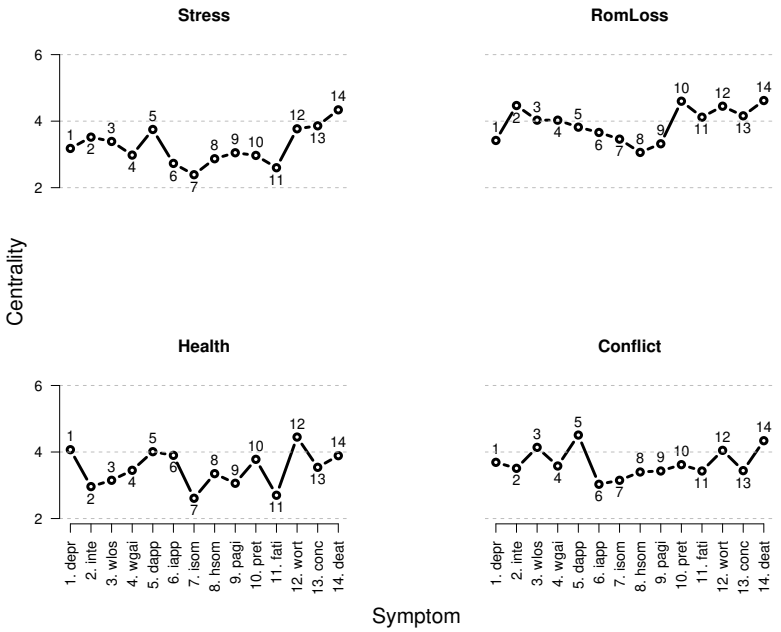

Figure 4.3: Centrality of symptoms in the four stressful life event groups. The top left panel depicts symptom centrality after stress (Stress); the top right panel after a romantic loss (RomLoss); the bottom left panel after health problems (Health); and the bottom right panel after an interpersonal conflict (Conflict). The x-axis represents the 14 disaggregated symptoms of an episode of MD while the y-axis represents centrality (defined as the sum of the tetrachoric correlations between a symptom and all the other symptoms in the network). *depr*: depressed mood; *inte*: loss of interest; *wlos*: weight loss; *wgai*: weight gain; *dapp*: decreased appetite; *iapp*: increased appetite; *isom*: insomnia; *hsom*: hypersomnia; *pagi*: psychomotor agitation; *pret*: psychomotor retardation; *fati*: fatigue; *wort*: feelings of worthlessness; *conc*: concentration problems; *deat*: thoughts of death.

Source of differences in correlation networks

This analysis contrasted the fit of model I to model II, as described in the Method section, and was based on a categorical one-factor model in which the following correlations were allowed to be estimated in all models, but constrained to be equal across SLE groups[3] ($\chi^2 = 499.15$, df = 72, RMSEA = 0.053, CFI = 0.863)[4]: weight loss with decreased appetite, psychomotor retardation with fatigue, insomnia with psychomotor agitation and hypersomnia with fatigue. Table 2 presents the results of fitting model I and model II to the data as well as a test of their relative fit. This test indicated that model II fitted the data significantly worse than model I ($p < 0.001$). This means that the differences in the correlation networks of the four SLE groups cannot be explained by underlying

[3] The choice for these correlations in particular was based on the modification indices in Mplus.

[4] The χ^2 statistic and CFI suggest poor fit but, given the low standardized residuals, we conclude nonetheless that the model fitted the data reasonably well.

differences in acute liability to develop a DE[5].

Table 4.2: Goodness-of-fit statistics and χ^2 difference tests for models testing factorial invariance of depressive symptoms across SLE groups

	χ^2	df	RMSEA	CFI	χ^2_{diff}	df_{diff}	p
Model I	769.97	303	0.054	0.856			
Model II	797.67	342	0.050	0.860	81.64	39	<0.001

SLE: stressful life event; df: degrees of freedom; RMSEA: root mean square error of approximation; CFI: comparative fit index; χ^2_{diff}: χ^2 statistic of the χ^2_{diff} difference test; df_{diff}: degrees of freedom of the χ^2 difference test; p: p value of the χ^2 difference test

Discussion

Two theories about the nature of psychiatric disorders—the common cause and the network hypothesis, respectively—postulate competing reasons for why SLEs influence depressive symptoms differently. The main goal of the present study was to investigate the predictions of these two hypotheses, an endeavor that is, to our knowledge, the first of its kind. To this end, we constructed networks of disaggregated depressive symptoms for four SLE groups based on participants with a DE. We compared these networks in a descriptive fashion, and assessed whether differences in the networks were (1) statistically significant and (2) best explained by underlying differences in acute liability to develop a DE. Our main results are that SLEs influence the correlations between depressive symptoms in markedly different ways; these differences are significant and cannot be explained by underlying differences in acute liability to develop a DE. That is, for example, the generally stronger correlations between depressive symptoms after a romantic breakup are not due to the fact that people have a higher liability to develop a DE after a romantic breakup compared with other SLEs.

Our results are not compatible with a common cause perspective. If a psychiatric disorder arises as predicted by the common cause hypothesis, then exogenous variables like SLEs should influence depressive symptoms only indirectly, via the common cause (i.e., acute liability to develop a DE; see panel A of Figure 4.1). And, as such, differences between SLEs in their impact on depressive symptoms should arise due to underlying differences in the common cause. A common cause model might be adjusted so that it fits in the SLE groups with equal strength of the associations between a DE and its symptoms; e.g., by allowing (1) SLEs to influence the symptoms directly or (2) some of the residual variance of symptoms to be correlated. However, such a model might fit but still violates the idea of a common cause through which exogenous variables (SLEs) exert their influence.

A network perspective on psychiatric disorders explains our results in a natural way: the symptoms and direct (causal) relations between them are the causes of a psychiatric disorder. As such, the network perspective predicts that exogenous variables (SLEs) will

[5] For the psychometrically interested reader: we have also fitted a strong factorial invariance model in which loadings and thresholds were constrained to be equal across groups (factor means and variances freely estimated in all but the first group, scaling factor fixed to 1 in all groups). Also in this case, the baseline model (in which loadings and thresholds were freely estimated across groups) was the preferred model.

make an impact on the symptoms directly without an intervening common cause: for example, a romantic breakup results in more hypersomnia than a health problem. We have not directly tested this hypothesis (i.e., we have not fitted a model like panel B in Figure 4.1 to the data, which is not possible given available methodological tools and the cross-sectional nature of the data), but, given that a common cause explanation is unlikely, the network perspective is currently the most plausible candidate for explaining the differences between SLEs in their influence on depressive symptoms.

If a network perspective on psychiatric disorders is accurate, then how do associations between symptoms arise? Suppose that depressed mood, insomnia, fatigue and concentration problems are strongly associated in someone: e.g., when Alice has trouble concentrating at work, she easily feels self-reproach for not being able to focus. Now, in such a strongly associated network of symptoms it theoretically takes only one symptom to become present as a results of an interpersonal conflict for instance—e.g., insomnia—for a syndrome to develop; for example, via the following sequence of events: insomnia → fatigue → concentration problems → self-reproach → depressed mood. It is likely that such connections between symptoms are governed by distinct pathological mechanisms. That is, a network perspective hypothesizes that symptoms will have partly distinct etiologies. For example, the connection between insomnia and fatigue will probably involve more physiological homeostatic mechanisms while the connection between depressed mood and self-reproach will be governed by more cognitive mechanisms. Moreover, individual differences are likely to arise in exactly these pathological mechanisms such that Alice will feel fatigued after only one sleep-deprived night while Bob can endure four sleepless nights without developing fatigue. Therefore, given that the analyses in this chapter indicate that such a network of connected symptoms might portray an accurate picture of DEs, we need an alternative research agenda that promotes the discovery and analysis of pathological mechanisms that govern individual symptoms and connections between them, currently not a focus in research into the etiology of psychiatric disorders.

A conceptual shift to a network perspective has clinical implications, for example the identification of people who are at risk of developing a DE, or its more severe counterpart, MD. We have shown that the centrality of certain symptoms in the correlation networks varies depending on the nature of the precipitating event; for example, loss of interest is a central symptom after a romantic breakup but relatively peripheral after a health problem. What might this difference in centrality imply in terms of risk for developing a DE? The centrality of a symptom could be interpreted as an indicator of how risky the presence of that symptom is for the development of a full-blown syndrome: a central symptom is one that is strongly connected (i.e., correlated) to the other symptoms in the network. As such, when someone develops such a symptom, there is a substantial risk that other symptoms will subsequently emerge as well, potentially resulting in a depressive syndrome. The present findings generate testable hypotheses with respect to which SLEs in combination with what symptom(s) might most likely result in a diagnosis of a DE in the future. One such hypothesis would be that people, after having experienced a romantic breakup, who present themselves with loss of interest have an elevated risk of developing a DE compared with those (1) with the same SLE but with other, peripheral, symptoms and (2) with the same symptom but with another SLE for which loss of interest is not a central symptom.

Our results have conceptual and philosophical implications regarding the nature of psychiatric disorders. The most common approach to understanding psychiatric disorders has been 'essentialism' (Kendler, 2006; Kendler & Baker, 2007): all important properties of a psychiatric disorder arise from a single causal process roughly analogous to the way in which all features of Down's syndrome arise from the presence of all or part of an extra 21^{st} chromosome. The common cause model is consistent with an essentialist model in

that all the symptoms of a DE arise from a common process, analogous to an essence. Alternatively, we argue that a different concept of the nature of psychiatric disorders, mechanistic property clusters (MPCs), may be a more accurate model (Kendler et al., 2011). This theory suggests that psychiatric disorders are more accurately defined in terms of mutually reinforcing networks of causal mechanisms. The network hypothesis is closely related to the concept of MPCs in suggesting that psychiatric disorders arise from interactions between their component symptoms rather than from some underlying essence. The findings presented in this chapter are, to our knowledge, the first empirical piece of evidence that such models might be accurate in their portrayal of psychiatric disorders.

Our findings should be interpreted within the context of some limitations. First, the participants were Caucasian twins born in the US state of Virginia. As such, we cannot be sure whether our findings will generalize to other populations. Second, the basic model that we used for comparing model I and model II did not fit the data well, and it might therefore be argued that the results of that analysis should be interpreted with some degree of caution. The most plausible reason for this lack of good fit was that our sample was a selection of only those people who reported at least two aggregated symptoms, thus not including participants with one symptom or no symptom at all (the model fitted the pre-selection sample data well). We repeated the analysis twice: (1) on the disaggregated data with a four-factor model that was not interpretable from a substantive point of view but fitted the data better; and (2) on the aggregated data for which a one-factor model fitted the data better than the model we reported on. In both cases, the outcome was identical to the one reported here. Third, with the available data, we cannot rule out the possibility that some of the covariation between some of the symptoms is due to an underlying, latent, mechanism (i.e., a common cause of some symptoms, rather than the common cause model as depicted in panel A of Figure 4.1). However, we note that for many of the depressive symptoms, direct relations appear to be more likely: e.g., that it is the actual experience of not sleeping that makes you tired (instead of a common underlying mechanism that causes both insomnia and fatigue). Finally, we investigated a limited range of all possible symptoms in the context of a limited number of stressors. Also, because the inclusion criterion for this study was less stringent than having a diagnosis of MD (i.e., two or more co-occurring aggregated depressive symptoms sufficed), the results paint a picture of DEs and, as such, cannot be straightforwardly generalized to their more severe counterpart, MD. That said, it must be noted that in this particular sample, 29% of participants did have a diagnosis of MD, a percentage that is almost three times higher than what is normally reported in MD studies.

Chapter 5

Major depression as a complex system

Abstract

What is major depression? In this paper, contrary to a latent variable perspective on psychopathology, we argue that major depression should be characterized as a complex dynamical system in which symptoms (e.g., insomnia and fatigue) causally interact with one another: insomnia → fatigue → psychomotor retardation. Next, we hypothesize that individual people can be characterized by their own network with unique architecture and resulting dynamics. With respect to architecture, we argue that individuals vulnerable to developing major depression are those with strong connections between symptoms: e.g., for a particular person, a sleepless night has a strong influence on feeling tired the next day. Such vulnerable networks, when pushed by external forces such as stressful life events, are more likely to end up in a depressed state (i.e., "depression" attractor); whereas networks with less strong connections tend to remain in or return to a healthy state (i.e., "health" attractor). We show this with a simulation in which we model the probability of a symptom becoming 'active' in a person as a logistic function of the activity of its neighboring symptoms. Additionally, we show that this model explains some well-known empirical phenomena (e.g., spontaneous recovery) and accommodates both continuous and taxonomic views on major depression. Finally, we elaborate on how therapeutic strategies (e.g., cognitive behavioral therapy) can be understood within this causal systems perspective. To our knowledge, we offer the first intra-individual, symptom-based, process model with the potential to explain the empirical reality of major depression.

Adapted from: Cramer, A. O. J., Giltay, E. J., van Borkulo, C. D., Kendler, K. S., van der Maas, H. L. J., Scheffer, M., & Borsboom, D. (*in preparation*). I feel sad therefore I don't sleep: Major depression as a complex dynamical system.

> "Slowly, over the years, the data will accumulate in your heart and mind, a computer program for total negativity will build into your system, making life feel more and more unbearable. But you won't even notice it coming on, thinking that it is somehow normal, something about getting older, about turning twelve or turning fifteen, and then one day you realize that your entire life is just awful, not worth living, a horror and a black blot on the white terrain of human existence. One morning you wake up afraid you are going to live."
>
> (Wurtzel, 2002)

Major depression (MD) imposes a heavy burden on people suffering from it. Not only are the symptoms of MD themselves debilitating, but their potential consequences such as job loss and facing stigmatization and rejection by other people can be equally detrimental to long-term physical and mental health (Greden, 2001; C. L. Hammen & Peters, 1978; Murray & Lopez, 1997; J. Wang, Fick, Adair, & Lai, 2007). Combined with the fact that major depression approximately affects 17% of the population at some point in their lives, denoting MD as one of the biggest mental health hazards of our time is hardly an overstatement (Kessler et al., 1994; Lopez, Mathers, Ezzati, Jamison, & Murray, 2006; WHO, n.d.). It should therefore not come as a surprise that vast amounts of time and money in clinical research have been allocated towards elucidating the causes of MD and effective ways to eliminate them.

One of the key questions that have to be asked first in order to investigate causes and design treatment interventions is what MD is; or, more generally, what a mental disorder is. The very notion of separate mental disorders, each associated with a specific set of symptoms, was first pioneered by Kraepelin (1923) and Lewis (1934) who, independently from each other, observed that particular symptoms tend to 'co-exist' with some but not all other symptoms. For example, depressed mood and feelings of worthlessness were seen together quite frequently in patients whereas depressed mood and disorganized thinking were not. Many such observations later culminated in the definition of distinct mental disorders, designating depressed mood and feelings of worthlessness as symptoms of MD and disorganized thinking as a symptom of schizophrenia. Put in statistical terms, the setup of the current classification system is based on the fact that some symptoms are more strongly correlated with one another (e.g., MD symptoms) than with others (e.g,. MD symptoms with symptoms of schizophrenia; see also C. A. Hartman et al., 2001). Now, the million-dollar question is why psychopathological symptoms show these particular correlational patterns. For many decades, the short answer to this question has been that mental disorders are latent variables, *common causes* of their symptoms, analogous to a lung tumor that causes shortness of breath and coughing up blood (see also Borsboom & Cramer, 2013). In this chapter, we present an alternative, namely that MD should be characterized as a complex dynamical system of interacting symptoms. First, we briefly review current conceptualizations of MD as a common cause of its symptoms. Next, we outline our view of MD and show that simulated networks have characteristics that are well-known in the empirical realm, for example spontaneous recovery; and accommodate both continuous and taxonomic views on major depression. Finally, we elaborate on how therapeutic strategies (e.g., cognitive behavioral therapy) can be understood within this causal systems perspective.

MD as a common cause of its symptoms

The general idea that symptoms of a mental disorder are attributable to the same cause (i.e., the *common cause view*) permeates—in explicit or implicit form—the field of psy-

chiatry and clinical psychology (e.g., psychosis is caused by hyperdopaminergia; Howes & Kapur, 2009) and is reflected in its mathematical formalization in psychometrics, the generic latent variable model (Borsboom et al., 2003; Cramer et al., 2010). In this model, the symptoms of MD cluster together because they share a common cause, MD: For example, the high correlation between insomnia and fatigue is hypothesized to be due to the fact that both were caused by the same underlying disorder (MD). In that sense, this model equates mental disorders such as MD to medical conditions, for example a lung tumor: in patients with such a tumor, symptoms such as chest pains, shortness of breath and coughing up blood are caused by the physical presence of the tumor. Likewise, the symptoms of Huntington's disease are caused by an abnormal length of the Huntingtin gene resulting in a mutant Huntingtin protein (Plomin, DeFries, McClearn, & McGuffin, 2008; Walker, 2007). Additionally, in both medical conditions, there is a clear-cut distinction between people with and without the disease: all people with Huntington's disease (or: lung cancer) have the mutant protein (or: tumor) and all people without Huntington's disease (or: lung cancer) do not have that protein (or: tumor).

We fitted a latent variable model—that is, the statistical equivalent of the common cause view as described above—to empirical data on the nine symptoms of MD—as they are indicated in DSM-IV (APA, 1994)—assessed for the previous year from the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPUD; Kendler & Prescott, 2006; Prescott et al., 2000). To be more precise, we fitted a one-factor model to this data and the results of this analysis are shown in Figure 5.1. The model fits the data nicely ($\chi^2 = 233.7, df = 27, p < .001$, CFI = .998, TLI = .997, RMSEA = .029) so one could argue: since it fits, why is the latent variable/common cause model not a good model to describe covariation among MD symptoms? One reason has to do with what happens when one fits a one-factor model to the exact same data but with the *disaggregated* instead of the *aggregated* MD symptoms. It is common practice in most, if not all, empirical papers that the latent variable model is fit onto the aggregated data in which symptoms such as insomnia and hypersomnia are collapsed (i.e., aggregated) into one symptom, sleep problems[1]. To us, these aggregations are suboptimal since, for example, there is good reason to believe that insomnia and hypersomnia have very different functions and might even be part of distinct subtypes of MD (more on this in the next section). Nonetheless, when a common cause on MD is accurate, then one should expect that the latent variable model would hold for the disaggregated symptoms as well. In the VATSPUD data, this is clearly not the case: the one-factor model fits poorly on the disaggregated symptom data ($\chi^2 = 3366.49, df = 77, p < .001$, CFI = .967, TLI = .961, RMSEA = .069). Naturally, it is possible to tweak the model in such a way that it does fit. In this example, the model fit for the disaggregated data was good ($\chi^2 = 432.55, df = 69, p < .001$, CFI = .996, TLI = .995, RMSEA = .024) when allowing seven residual variances—based on modification indices greater than 100—to be correlated[2]. However, although this *statistically* tweaked model fits disaggregated data, it is, *theoretically* speaking, a far cry from a common cause model: for how could one claim that a certain variable (MD in this case) is the common cause of a set of observable symptoms while at the same time allowing that, say, weight loss and decreased appetite are (cor)related?

It should be stressed that a pivotal consequence of adhering to a common cause view

[1] Other aggregated symptoms are psychomotor disturbances (agitation and retardation) and weight problems (weight loss and gain, increased and decreased appetite.

[2] The following residual variances were allowed to be correlated: weight loss with weight gain, weight loss with decreased appetite, weight loss with increased appetite, weight gain with decreased appetite, weight gain with increased appetite, decreased appetite with increased appetite, insomnia with psychomotor agitation, and psychomotor retardation and fatigue

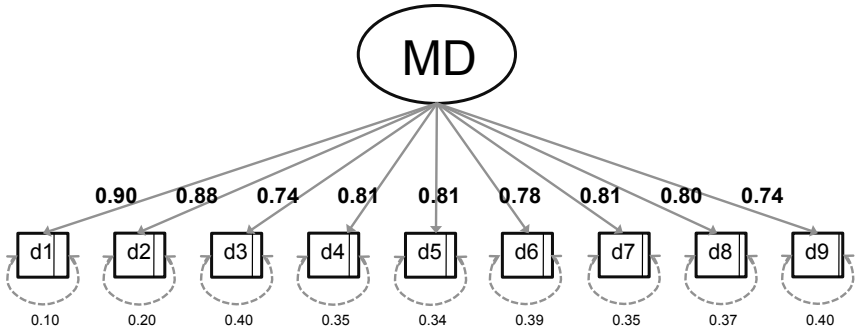

Figure 5.1: A one-factor model of major depression based on data from the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATPSUD). The circle on top of the figure represents the latent variable "MD" whose metric was fixed such that its mean was 0 and its variance 1. The nine squares represent the nine aggregated symptoms of MD according to DSM-III-R. The grey lines (i.e., edges) from the latent variable "MD" to the squares represent the factor loadings. The vertical line in each of the squares represents the threshold above which a "yes" response is given: the further to the right the line, the higher the threshold. The dashed lines going in and out of the squares represent the residual variance. *d1*: depressed mood; *d2*: loss of interest; *d3*: weight problems; *d4*: sleep problems; *d5*: psychomotor problems; *d6*: fatigue; *d7*: feelings of worthlessness; *d8*: concentration problems; *d9*: thoughts of death.

is that one assumes that there are no causal relations between symptoms (i.e., no arrows between symptoms in Figure 5.1). From the common cause view, insomnia and fatigue 'co-exist' because they are caused by MD and *not* because they are causally related: e.g., insomnia → fatigue. In latent variable models—again, the model of choice when adhering to a common cause view—the "local independence" axiom translates this idea into statistical terms: conditioning on the latent variable (i.e., MD in Figure 5.1 renders the symptoms statistically independent (e.g., Holland & Rosenbaum, 1986; Lord, 1953; McDonald, 1981). Again, this idea makes sense for quite a number of medical diseases: involuntary jerking and lack of impulse control are likely not causally related but both caused by the mutant Huntingtin protein.

The pervasiveness of the common cause view has its roots in a mix of the philosophical traditions of *disease realism*, *reductionism* and *essentialism* (e.g., Haslam, 2000; Haslam & Ernst, 2002; Lilienfeld & Marino, 1999; R. Kendell & Jablensky, 2003; Thornton, 2000). Briefly, the combined views claimed mental disorders are *real* in the sense that it has "a real, substantial existence regardless of social norms and values, and exist independent of whether they are discovered, named, recognized, classified or diagnosed" (Freitas, 2007). As such, a mental disorder like MD is ultimately grounded in tractable biological abnormalities and a true distinction exists between people with and without MD. While we know that this is potentially true for a number of medical conditions, there is no convincing evidence for the validity of this conceptualization of MD. There is no clear-cut distinction between people with and without MD: there exists no set of biological abnormalities that are always present in people with MD vis-a-vis people without MD. Even worse, after decades of intensive research, no one has been able to even come up with a likely candidate to function as the common cause of MD (research into common causes for other clusters of psychopathological symptoms share the same fate). One could argue that we have not found those common causes *yet*: we will find the root causes of mental disorders like MD some day but we simply need better tools to discover them (Bollen, 2002; Borsboom, 2008b). A different take on the matter is simple:

there are *no* common causes to be found because they do not exist (Cramer et al., 2010; Cramer, van der Sluis, et al., 2012).

MD as a network of interacting symptoms

So what, then, is responsible for the established fact that symptoms of MD cluster together (i.e., are highly correlated)? Our alternative conceptualization of MD starts out with the exact opposite assumption of a common cause view concerning relations between symptoms: MD is a network of symptoms that *do* causally interact with one another. That is, insomnia, fatigue and concentration problems are not highly correlated because they are caused by MD but because they are causally related to each other: insomnia → fatigue → psychomotor retardation (Borsboom, 2008b; Cramer et al., 2010; Cramer, Borsboom, Aggen, & Kendler, 2012). Such direct relations make more sense than postulating a common cause: why would one need a common cause to explain why not sleeping and feeling tired are highly correlated? Likewise, if one considers the symptoms "depressed mood" and "self-reproach" it is plausible to assume a direct causal relationship: if one suffers from depressed mood long enough, at least a subset of the people experiencing this might start developing feelings of self-reproach because of these depressed feelings; and, again, this makes more sense than postulating a common cause for whose existence we have no evidence. Thus, both the common cause as well as the *network perspective* on psychopathology share the starting point that some symptoms are more strongly associated with one another than with others; they differ, however, in the hypothesized reason for these association patterns: a common cause, from a common cause perspective, and direct relations between these symptoms, according to the network perspective.

There are several ways to construct a network for any given set of variables. Ordinary correlations are a potential starting point but, as we have emphasized elsewhere (Cramer, van der Sluis, et al., 2012), their usefulness is limited since a high correlation need not be indicative of a direct relation between two variables. For example, it might be the case that two nodes in a network—say, self-reproach and fatigue—are highly correlated but only do so because they are both caused by a third node in the network, say, insomnia (Borsboom & Cramer, 2013). A fruitful alternative to determine which nodes are connected in a network is by exploring conditional independence relations. The package *pcalg* for R (Kalisch, Maechler, Colombo, Maathuis, & Buehlmann, 2012) can be used for the discovery of conditional independence relations from observed data. In a nutshell, the package starts out with a graph in which all nodes are connected to one another without directionality. Next, it determines conditional independence relations because these imply which nodes should be connected and which not. For example, consider the network in panel B of Figure 5.2, suppose that this network is the 'true' network that gave rise to a dataset: thus, in this example, both variables X and Y are caused by Z (i.e, a *common cause* structure). Now, such a network implies that X and Y are conditionally independent given Z (Pearl, 2000). Generally, two variables X and Y are said to be conditionally independent given a third variable Z, denoted as $X \perp Y|Z$, iff $\Pr(Y|X,Z) = \Pr(Y|Z)$. That is, X and Y are conditionally independent if the probability of Y given X and Z is the same as the probability of Y given Z alone: values of X thus give no additional information about the probability of Y occurring. And because X does not contribute information, beyond the information contributed by Z, to the probability of Y, the program removes the edge between X and Y. *pcalg* searches the dataset for all such relations (for all possible combinations of nodes in the network including all interactions) to delete edges from the completely connected network. In the next step, it infers the directionality of the edges by, for example, exploiting the special

case of the *collider* (see panel C of 5.2). A collider, Z in panel C of Figure 5.2, is a node that is the outcome of two other variables, X and Y. In this particular case, X and Y are conditionally *dependent* given Z. Once search algorithms like *pcalg* find such colliders, the edges are oriented accordingly. On a cautionary note, a limitation of these and similar search algorithms is that a set of conditional independencies can be consistent with multiple graphs (these graphs are said to be *Markov equivalent*): for example, the networks in panel A (i.e., *causal chain*) and B (i.e., common cause) of Figure 5.2 imply the same, single, conditional independency: $X \perp Y|Z$. This means that when using such search procedures, it would be premature to claim the discovery of *the* network that gave rise to the conditional independence relations between a given set of variables. A network for a particular sample discovered through causal discovery algorithms such as *pcalg* would need independent verification in unrelated samples; and/or, additional variables could be added as extra nodes in the network in order to facilitate the process of determining directionality of the edges.

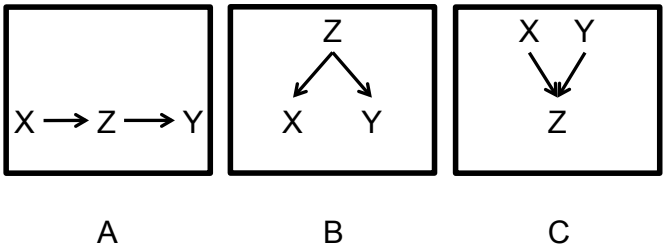

Figure 5.2: An illustration of the three most important causal relations that can be discovered through tracking conditional independence relations. Panel A shows a chain structure: Z functions as a mediator between X and Y. Panel B shows a common cause structure: Z acts as the common cause of both X and Y. Panel C shows a collider structure: Z is a common effect of both X and Y.

Figure 5.3 presents such a graph derived from *pcalg* for the VATSPUD data. In this figure, each symptom is represented as a node while each line (i.e., edge) between two nodes implies that these nodes are conditionally dependent, given all combinations of other nodes in the network. We have omitted arrows in this figure so all edges are interpreted as bidirectional connections. The positioning of the nodes is such that nodes with many connections with other nodes are placed towards the middle of the graph; while nodes with relatively few connections with other nodes are placed towards the periphery of the graph. The first thing that stands out when inspecting Figure 5.3 is that a couple of symptoms appear to be isolated from the remainder of the network: most noteworthy in this respect are weight loss (*wlos*), weight gain (*wgai*), increased appetite (*iapp*) and hypersomnia (*hsom*). Pertaining to the latter three, the fact that these appear to be somewhat isolated from the other symptoms of MD might not come as a surprise since the distinction between MD with typical (e.g., weight loss, decreased appetite, insomnia) versus atypical (e.g., weight gain, increased appetite, hypersomnia) symptoms was first observed and articulated into theory many decades ago (e.g., D. F. Klein & Davis, 1969) and is currently recognized in the most recent version of the DSM (i.e., DSM-IV; APA, 1994). Second, some symptoms are clearly more central than others: these nodes have relatively many connections with other nodes in the network. Consistent with the diagnostic importance that DSM-IV assigns to depressed mood and loss of interest—either one of these symptoms must be present for a diagnosis of MD—these symptoms (*depr* and *inte*) rank among the most central symptoms in the MD network. On a final note, many of the connections in Figure 5.3 make intuitive sense: for example, the connec-

tion between insomnia (*isom*) and fatigue (*fati*); the connection between fatigue (*fati*) and concentration problems (*conc*); and the connection between feelings of worthlessness (*wort*) and thoughts of death (*deat*).

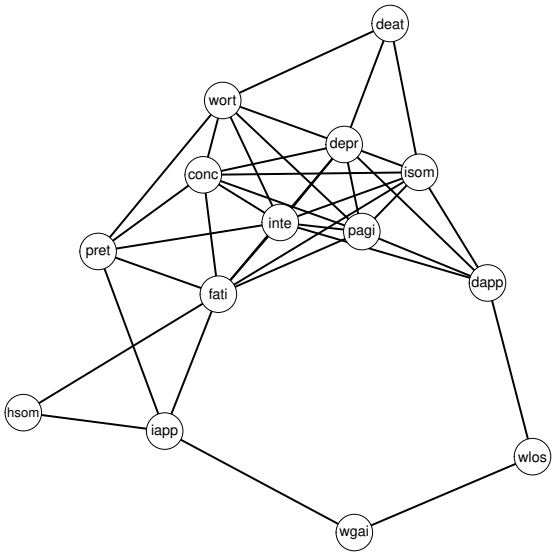

Figure 5.3: A network model of major depression derived from conditional independence relations and based on the VATSPUD data. Each node in the figure represents one of the 14 disaggregated symptoms of MD according to DSM-III-R. A line (i.e., edge) between any two nodes means that they are conditionally dependent, given all possible subsets of other nodes in the network. The absence of an edge indicates conditional independence. *depr*: depressed mood; *inte*: loss of interest; *wlos*: weight loss; *wgai*: weight gain; *dapp*: decreased appetite; *iapp*: increased appetite; *isom*: insomnia; *hsom*: hypersomnia; *pret*: psychomotor retardation; *pagi*: psychomotor agitation; *fati*: fatigue; *wort*: feelings of worthlessness; *conc*: concentration problems; *deat*: thoughts of death.

Recent (circumstantial) evidence appears to be in favor of the network model for MD (Cramer, Borsboom, et al., 2012). For example, the death of a spouse triggers depressed mood (*depr*) and loss of interest (*inte*) while health problems trigger insomnia (*isom*) and psychomotor retardation (*pret*) in Figure 5.4 (i.e., the solid arrows); and this phenomenon cannot be explained by underlying differences in the common cause (i.e., the dashed arrows in Figure 5.4): for example, the fact that the death of a spouse triggers more depressed mood and loss of interest compared to health problems, cannot be explained by the death of a spouse causing a higher score on the common cause/latent variable "MD" compared to health problems. This result points to the unique role that MD symptoms appear to play in the pathogenesis of MD: each receives, potentially unique, input from external variables such as stressful life events. And this unique role of individual symptoms is hard, if not impossible, to explain from a latent variable or common cause perspective where, by the very implication of positing a common cause, all external influences (such as stressful life events) should run via the common cause.

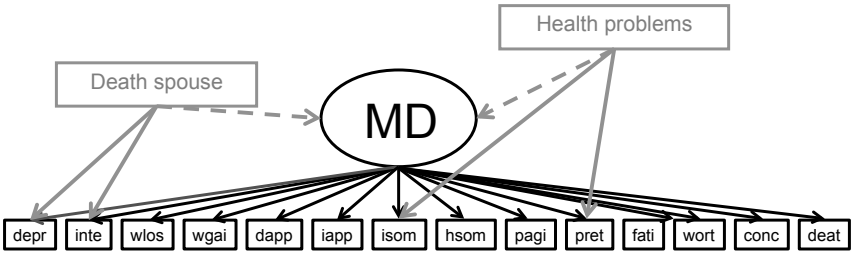

Figure 5.4: Two potential ways in which stressful life events (i.e., the death of a spouse and health problems) can impact major depression. The circle with "MD" represents the latent variable "major depression" while the rectangular boxes represent the 14 disaggregated symptoms of MD. First, a stressful life event can influence the latent variable MD directly and not the symptoms (i.e., dashed lines from *Death spouse* and *Health problems* to *MD*). This setup is consistent with a latent variable perspective. Second, a stressful life event can influence the symptoms directly and not the latent variable (i.e., solid lines from *Death spouse* and *Health problems* to *depr*, *inte*, *isom* and *pret*). This setup is consistent with a network perspective. *depr*: depressed mood; *inte*: loss of interest; *wlos*: weight loss; *wgai*: weight gain; *dapp*: decreased appetite; *iapp*: increased appetite; *isom*: insomnia; *hsom*: hypersomnia; *pret*: psychomotor retardation; *pagi*: psychomotor agitation; *fati*: fatigue; *wort*: feelings of worthlessness; *conc*: concentration problems; *deat*: thoughts of death.

The intra-individual network model of MD

The network in Figure 5.3 is based on inter-individual data and as such, without additional testing of the mapping between what happens at the inter-individual level and what happens within an individual, could not serve as an intra-individual network model of MD. How, then, could we model what happens at the intra-individual level? Let us start with explication how one could conceive such a model. First, we assume that symptoms can be 'on' (1) or 'off' (0) and can influence one another in time if they are connected in the graph based on conditional independencies derived from the VATSPUD data (Figure 5.3). Symptoms that are connected in this graph are called *neighbors* of one another. In Figure 5.3 for example, hypersomnia has only two neighbors (increased appetite and fatigue) and only these neighbors can influence hypersomnia. Likewise, psychomotor retardation has five neighbors (loss of interest, increased appetite, fatigue, feelings of worthlessness and concentration problems). The *state* (0: 'on' or 1: 'off') of a symptom i at time t is denoted X_i^t. The connection between symptom i and j is denoted W_{ij} and can be considered as a binary variable (connection is either present (1) or absent (0) as in the connections in Figure 5.3) or a continuous variable (for example ranging from -1 to 1 reflecting partial correlations). Furthermore, there is no autocatalysis: the weights W_{ii}—that is, a connection (self-loop) between a symptom and itself—are set to 0. The matrix $\mathbf{W}$ contains all weights of the connections between $J = 14$ symptoms.

Next, we assume that the probability of a symptom to turn 'on' (i.e., becoming present/active in a 'person') depends monotonically on the activation of its neighbors: the more neighbors of symptom i are 'on' at a given point in time, the higher the probability that symptom i, at a later point in time, becomes present itself. The *total activation function* of symptom i at time t is:

$$A_i^t = \sum_{j=1}^{J} W_{ij} X_j^t \tag{5.1}$$

Thus, when connections between symptoms are stronger and more symptoms are

turned 'on', A becomes increasingly large. As a next step, the probability of a symptom i becoming active at time $t+1$ depends on A_i^t and is expressed as follows:

$$P(X_i^{t+1} = 1) = \frac{1}{1 + e^{a_i(b_i - A_i^t)}} \tag{5.2}$$

Note the similarity of this probability function with the probability function for an item in a Rasch model (Rasch, 1960) and the conditional probability of a variable given its neighbors in the classical Ising model (Ravikumar, Wainwright, & Lafferty, 2010) (see Appendix F for more details on the exact correspondence of our model with the Rasch model and the Ising model). In the current function, the parameter b_i gives the threshold of symptom i. Symptoms with higher thresholds require more activation, while symptoms with lower thresholds are easily activated. A special case arises when A is equal to b_i, that is, when the amount of activation of the neighbors of symptom i is exactly equal to the threshold of symptom i. In that case, the probability of symptom i becoming active is exactly 1/2. Given known prevalence differences in MD symptomatology—for example: sleeping problems are more prevalent than thinking of suicide—threshold differences between symptoms appear to be a reasonable modeling assumption. The parameter a_i is a symptom-specific parameter that controls the steepness of the probability function: for higher levels of a_i, a given change in A_i^t results in a steeper change in the probability of symptom i becoming active. Together, formulas 5.3 and 5.2 control the dynamic behavior of the system. We implemented the model in the freely available program NetLogo (van Borkulo, Borsboom, Nivard, & Cramer, 2011)[3].

The above formulas describe the dynamic behavior of one network over time. Suppose we would simulate many of these networks to represent individuals, how might these networks differ in their architecture? We hypothesize two things: First, individuals' networks differ in terms of the connections between symptoms; that is, individuals might have a different W_{ij} matrix. In substantive terms for binary weights, this might mean, for example, that insomnia has no neighbors in Alice's network (i.e., for Alice, poor quality of sleep does not influence other symptoms in her network); while in Bob's network, insomnia has eight neighbors (i.e., in Bob's case, poor quality of sleep influences many other symptoms in his network). When the connections are considered to be continuous, then a strong connection between, say, insomnia and fatigue in someone's network of MD means that poor sleep one day has a substantial impact on feeling fatigued the next day: this person likely will feel tired after a night with poor sleep. On the other hand, a weak connection between insomnia and fatigue means that poor sleep one day has only a limited impact on feeling fatigued the next day: this person likely will not feel tired after a night with poor sleep. Second, we hypothesize that someone who is *vulnerable* to develop an episode of MD has a network in which the MD symptoms are generally strongly connected: according to formulas 5.3 and 5.2, if one symptom is developed in networks with high connectivity, then the probability of other symptoms quickly becoming activated as well is high, due to (1) the presence of relatively many connections (in case of W_{ij} containing binary weights) or (2) the overall strong connections (in case of W_{ij} containing continuous weights) in that network. The worst case scenario, in terms of vulnerability, is the combination of strong connections between symptoms and relatively low thresholds: in that particular case, not much activation of neighboring symptoms is needed to exceed the threshold of a symptom and, as such, activate it. On the other hand, relatively high thresholds might 'protect' a person from harm because in that case, despite strong connections, al lot of neighboring symptoms need to be activated in order to exceed the threshold. *Resilience* can be thought of as a network in which the symptoms

[3] The simulation tool can be downloaded at http://ccl.northwestern.edu/netlogo/models/community/Symptom%20Spread%20Model

are weakly connected: the development of one symptom is not likely to set off a cascade of symptom development culminating in an episode of MD; because the relations are not strong (or not omnipresent in the case of binary weights), thereby having a relatively small impact on the probability of symptom development.

To incorporate the possibility of different weight matrices and study the potential impact on the behavior of the network, we adapted the formula in 5.2 to include a connectivity parameter c, a number that is identical for all symptoms with which matrix W_{ij} is multiplied: the higher c, the more strongly the symptoms are connected:

$$A_i^t = \sum_{j=1}^{J} cW_{ij}X_j^t \tag{5.3}$$

For our simulations, c took on three values to create networks with low (lowest value of c), medium and high connectivity (highest value of c). At all time points during simulations, we tracked the global mood state, M, of these networks by computing the total number of activated symptoms at these time points. We simulated 10000 time points with this basic intra-individual MD model for each of the three values of c to investigate whether, as we predict, stronger connectivity results in higher levels of M. The a_i and b_i parameters for the probability functions of the 14 symptoms were derived from fitting a logistic regression to the VATSPUD data, in which each symptom was regressed on the total score of its neighbors in the network model as it is presented in Figure 5.3[4]. Detailed results as well as R scripts are available at www.aojcramer.com. The most important result is that, as we predicted, the higher the connectivity of the network, the higher the mood state M averaged over the 10000 simulations. That is, more strongly connected networks are more vulnerable in that they become "depressed" more easily.

Another important result has to do with the weakly connected network in particular. It is a well-known, but relatively understudied, phenomenon that quite some individuals who suffer from an episode of MD recover independently of treatment. Estimates of this *spontaneous recovery* range somewhere between 23% and 98% (e.g., see Krøgsboll, Hróbjartsson, & Gøtzsche, 2009; Kendler, Walters, & Kessler, 1997; Whiteford et al., 2012) depending on the exact time frame within which people with MD are observed and on whether these people have received treatment: for example, the spontaneous recovery rate is around 23% for untreated individuals within three months while it is around 90% for a community sample within a year combining both treated and untreated individuals. In treatment studies, these people are, when identified, omitted from the statistical analyses and preferably, the baseline period before the start of a treatment intervention is increased such that as little spontaneous recovery as possible 'contaminates' the results. While a sensible requirement when the aim is to study the effects of a treatment intervention, it has resulted in very limited knowledge about the naturalistic course of major depression (see: Posternak et al., 2006). As a result, there are no solid theories about how spontaneous recovery might come about. To our knowledge, the results of this basic simulation study are the first to hint at a possible mechanism through which spontaneous recovery might occur. See Figure 5.5 in which the mood state M (y-axis)

[4] Parameter estimates are available at www.aojcramer.com. Please note that the intercept (b_i) and slope parameters (c_i) derived from logistic regression for each symptom on the total score of its neighbors (S_i) (i.e., $P(X_i^{t+1} = 1) = \frac{e^{b_i + c_i S_i}}{1+e^{b_i + c_i S_i}}$); do not have a one-to-one mapping with our a_i and b_i parameters in $P(X_i^{t+1} = 1) = \frac{1}{1+e^{a_i(b_i - A_i^t)}}$. Such a mapping can be achieved by setting our a_i parameter to be equal to the c_i parameter of the logistic regression; and our b_i parameter to be equal to $-\frac{b_i}{c_i}$.

for a weakly connected network is shown for the first 2000 time points. The figure clearly shows spontaneous recovery: that is, there are points (e.g., at time points 50 and 600) where many symptoms are activated and, without any change to its parameters, the system recovers spontaneously to a state in which no symptoms are activated.

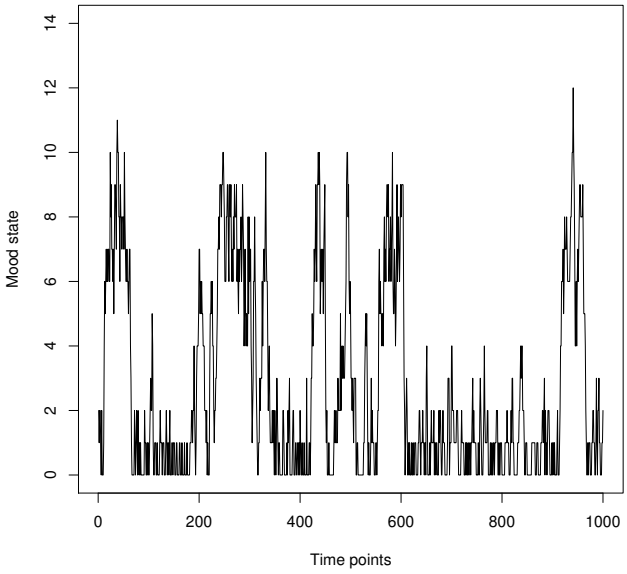

Figure 5.5: The mood state of the MD system over time. The x-axis represents the first 1000 time points of the basic simulation with the MD network model while the y-axis depicts the mood state; that is, the total number of activated symptoms at a given time point. The figure shows the presence of spontaneous recovery: there are distinct peaks in the mood state, which, without intervening in the parameters of the system, spontaneously return to lower mood states.

Vulnerability to develop MD: a catastrophe?

Our current knowledge of how stress (in interaction with vulnerability) can cause episodes of MD is based on inter-individual differences: for example, in Chapter 4, inferences about the impact of stress on symptoms of MD were made by comparing groups of individuals. As such, we have no clear idea of what happens inside an individual person when put under stress, which might or might not result in an episode of MD. Therefore, in order to investigate the impact of stress on intra-individual networks, we extend our basic network model of MD, explicated in the previous section, with a stress parameter S_i^t, a number for each symptom i that is added to the activation of the neighbors of i: the higher S_i^t, the more stress, the higher the total activation function, and thus the higher the probability that symptom i will become active. This results in the following modified activation function:

$$A_i^t = \sum_{j=1}^{J} W_{ij} X_j^t + S_i^t \tag{5.4}$$

Thus, in our model, stress S_i^t triggers the development of one or more symptoms of MD. It then depends on the strength of the connections between the MD symptoms ($\mathbf{W}$) whether the symptoms that were developed due to the stressor cause the development of other symptoms. If someone is vulnerable (i.e., symptoms are strongly connected), a mild stressor triggering the development of one symptom could be enough to trigger a cascade of symptom development eventually culminating in a full-blown episode of MD (i.e., large M). In a resilient person/network, a relatively large stressor might trigger the development of one or a few symptoms, but these symptoms will not likely cause the development of other symptoms. Consider an analogy with a bowl containing a ball (see also Figure 5.6): the ball is the MD network and the lowest point of the bowl is one of two possible mood states (alternatively called: *attractors* or *basins of attraction*), a *healthy* euthymic state and a *depressed* state. Now, when someone is resilient (i.e., at low risk to develop MD), stressors are capable of pushing the ball out of the healthy attractor but when the stressor diminishes, the ball quickly rolls back towards that healthy attractor. That is, in a resilient person, there is a quick homeostatic return to the prior mood state. Box A of Figure 5.7 illustrates this: larger stressors only slightly affect mood along the equilibrium attractor curve.

We hypothesize that the MD symptom networks of people who are vulnerable to develop MD (i.e., strongly connected networks) behave like what is known in the complex systems literature as a *cusp catastrophe model* (Ehlers, 1995; Flay, 1978; Goldbeter, 2011; M. T. Hubert, Braun, & Krieg, 1999; Thom, 1975; Zeeman, 1977). The catastrophe model is a mathematical model that can explain why small changes in some parameter (in our case: small increment in the stress parameter S_i^t) can, under certain circumstances, result in catastrophic changes in the state of a system (in our case: a catastrophic shift from a healthy to a depressed state, and vice versa). In the case of someone who is vulnerable to develop MD, even a small perturbation like a particularly bad day at work has the potential to push the ball outside the healthy attractor, causing it to roll towards another attractor, namely that of the depressed state (see also Figure 5.6). If this depressed basin of attraction is large and deep, the network of MD symptoms may become solidly locked in the depressed state. But under what circumstances can such a catastrophic shift from health to MD happen?

Box C of Figure 5.7 illustrates three features of a cusp catastrophe model that, combined, can bring about catastrophic semi-permanent changes in response to small perturbations. First, the equilibrium curves are folded backwards with multiple equilibria for a given stressor as a result: e.g., a divorce might result in the MD network ending in the healthy or depressed attractor. Second, the equilibrium attractor curve contains two *tipping points* that represent the border between the basins of attraction of the two alternative stable states on the upper and lower branches (Scheffer et al., 2009). Near such tipping points the equilibrium is *unstable* (dashed middle sections in Box C of Figure 5.7), meaning that modest disturbances (e.g., a small argument with a spouse over which restaurant to go to) may already move the network away from the healthy attractor instead of returning to it; and may even result in a large catastrophic affective shift to the depressed state. Third, the two alternative stable states of the MD are stable, implying that once the MD network has gone through a catastrophic affective shift, it tends to remain in that new affective state until the external input (i.e., stress) is changed back to a much lower level than was needed to trigger the episode of MD; a phenomenon known as *hysteresis* (see also Box C of Figure 5.7). That is, this model predicts that Bob, who developed an episode of MD in response to severe marital problems, will not recover automatically when, for some reason, the current marital problems are solved: more is needed, for example additional treatment, to trigger Bob's recovery.

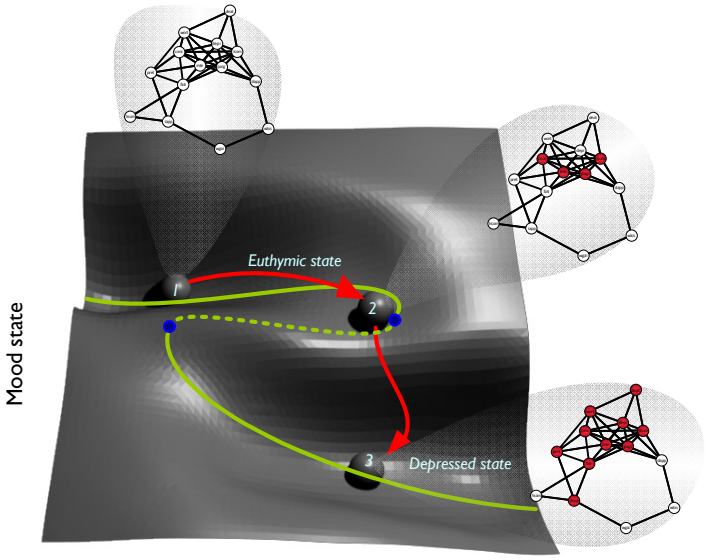

Figure 5.6: Three-dimensional landscape model for major depression (MD). Stable affective states are represented as curvatures (i.e., wells), in a two-dimensional surface. The model contains two tipping points, represented by blue dots, resulting in two equilibria or stable states (solid green lines in the two wells) that exist for a given stressor. The green dotted line marks the unstable state, as the equilibrium curve is folded backwards. The balls represent the MD network. Far from the bifurcation point (1), resilience is large because the basin of attraction in the deep well of the healthy state is larger with a high rate of recovery from perturbations. This state is typical for a MD network with weakly connected symptoms. Under gradual influence of stress, the well becomes shallower and the basin of attraction shrinks (ball progressing along the red arrow from 1 to 2). More symptoms become activated (marked by red nodes). The basin finally vanishes at a critical threshold (i.e., bifurcation points or tipping points), causing the ball to abruptly roll towards the depressed state attractor (progressing along the red arrow from 2 to 3).

The simulation

For the actual simulation that tested our hypotheses regarding the behavior of weakly, medium, and strongly connected MD networks in response to stress, we—similar to the basic MD network model—simulated 10000 time points for each of the three connectivity c values. For these three types of networks, we observed the impact of variation in the stress parameter: over the course of the 10000 time points S_i^t was repeatedly gradually increased as well as gradually decreased from -15 to 15 with small steps of 0.01 (for details of the simulation see www.aojcramer.com). The impact of the stress parameter on the behavior of the network was quantified by computing the average global mood state M: that is, since all stress parameter values were used multiple times during the simulation, we averaged mood states within 0.20 ranges of the stress parameter values.

Figure 5.8 shows the main results of the simulation: the x-axis represents stress, S_i^t, while the y-axis represents the global mood state M. The red line represents the mean number of activated symptoms (for stress parameter values within 0.20 ranges) when stress was increasing and the blue line represents the mean number of activated

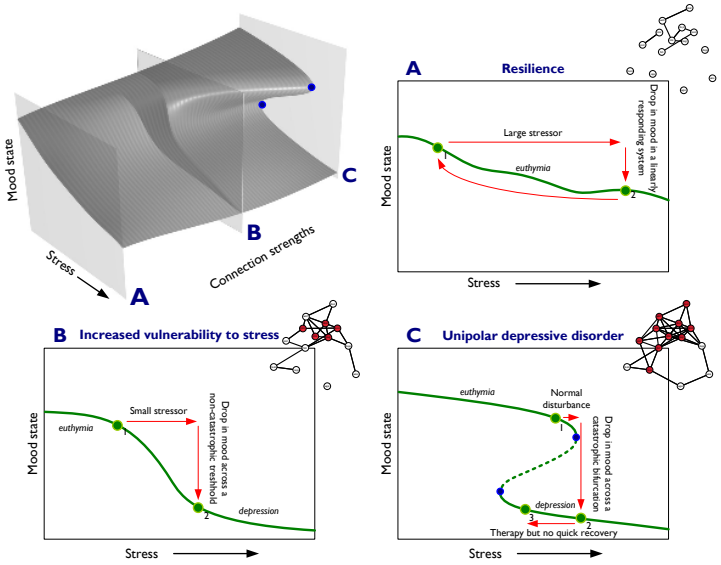

Figure 5.7: Diagrams of potential (critical) transitions in affective states in a cusp catastrophe model. The two control variables are connectivity of the MD network and stress. The equilibrium state of the mood system can respond in different ways to stress. Red arrows in the plots indicate the direction in which the system moves under the influence of stress. Box A: In resilient people, severe stress results in relatively modest changes in the affective state, the system quickly returns to its prior state. A resilient person is not immune to stress, but is able to re-establish equilibrium following stressful experiences. Box B: In more vulnerable people, relatively small stressors may induce a larger drop in the affective state. Small stressors thus can cause relatively large changes in the absence of true bifurcations, provided that the system is very sensitive along a certain range of stressors. Box C: The equilibrium is folded backwards in a model for some of the more severe forms of depressive disorders, resulting in two tipping points (i.e., critical thresholds or bifurcation points; marked by blue dots) and two alternative stable states for a certain set of values of the stress parameter. This represents the cusp catastrophe model. When the MD system is near a tipping point (i.e., critical threshold), a small disturbance may result in a catastrophic affective shift towards an episode of MD as the system jumps to the far-away attracting lower branch.

symptoms (for stress parameter values within 0.20 ranges) when stress was decreasing. As a general result, differences in the strength of connections between MD symptoms resulted in markedly different responses to external activation by stress. MD networks with low connectivity proved resilient (left panel of Figure 5.8): stress increments led to a higher number of developed symptoms in a smooth continuous fashion, and stress reduction resulted in a smooth continuous decline of symptom activation. The dynamics were very different for the networks with medium and strong connectivity, which proved vulnerable (middle and right panel of Figure 5.8, respectively). Here, the shift between healthy and depressed states was of a non-linear character as two tipping points appeared: a small increase in stress could lead to a disproportional reaction, resulting in a depressed state. Additionally, these vulnerable networks needed a significant decrease in stress to move the network back to a healthy state. The middle (medium connectivity) and right (high connectivity) panel of Figure 5.8 also clearly show that during the transition from the healthy to the depressed state or vice versa, a 'forbidden' zone (from 1 to 5

activated symptoms) was crossed that does not seem to function as a stable affective state. Moreover, hysteresis, as we anticipated earlier, was evident in the networks with strong connectivity: in order to make these networks return to a healthy state, the stress level needed to be reduced to a level far below the level of the initial external activation by stress. Thus, a cusp catastrophe was present for such networks. Thus, in sum, the results of this simulation as visualized in Figure 5.8 highly resemble the hypothesized behavior of MD networks as visualized in Figure 5.7. In the remainder of this section, we describe some of these results, and, if applicable, their connection with the empirical reality of MD, in greater detail.

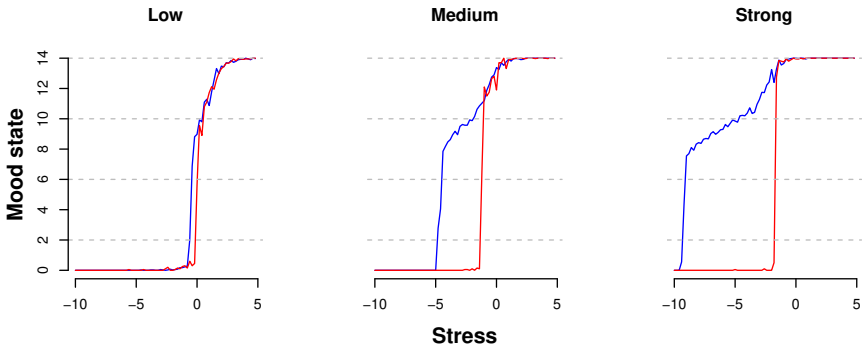

Figure 5.8: The mood state of the MD system in response to stress. The x-axis represents stress while the y-axis depicts the average mood state: that is, the total number of activated symptoms averaged over every 0.20 range of the stress parameter value). The red line depicts the situation where stress is increasing whereas the blue line depicts the situation where stress is decreasing. The three graphs represent, from left to right, the simulation results for networks with low, medium, and high connectivity, respectively.

Hysteresis

As we hypothesized, the results of the simulation of a network with high connectivity showed clear hysteresis: the amount of stress reduction needed to get the system into healthier mood states (i.e., only a few symptoms present) exceeds the amount of stress that tipped the system into depressive mood states in the first place. We do not know of other (simulation) studies that showed this phenomenon to be present in 'individuals' vulnerable to developing episodes of MD. It does seem to resonate with clinical observations concerning the non-linear course of affective shifts between healthy and depressed states that is frequently encountered in the empirical literature: it takes relatively little time to induce an episode of MD by mild stress, but it takes more time to recover from such an episode because more is needed for recovery, besides removing the stressor that initially triggered the development of the episode (e.g., Penninx et al., 2011).

One could argue, since we have used one specific setup of the simulation with specific choices for (empirically derived) parameters, that the hysteresis effect is an effect of this specific setup; and that, as such, the observed hysteresis is not a robust phenomenon that can be consistently associated with high connectivity networks. In order to check the robustness of the hysteresis effect, we repeated the simulations in which we systematically varied five parameters: (1) weights W_{ij}: either discrete (as in our original setup) or continuous (with uniformly distributed values between 0 and 1), (2) connectivity pa-

rameter c ranging from 1 to 3, (3) number of nodes J in the network: 5, 10, 15 or 20, (4) the b_i parameter which was randomly sampled from a normal distribution in three configurations: $\mu = 1$ and $\sigma = 0.25$, $\mu = 1.50$ and $\sigma = 0.50$ or $\mu = 2$ and $\sigma = 0.75$ and (5) the a_i parameter which was randomly sampled from a normal distribution in three configurations: $\mu = 1$ and $\sigma = 0.25$, $\mu = 1.50$ and $\sigma = 0.50$ or $\mu = 2$ and $\sigma = 0.75$. Each simulation consisted of 10000 time points and each combination of parameter values was simulated 20 times. For each simulation we computed a hysteresis effect H: for both increasing and decreasing stress (the same red and blue lines, respectively, as in Figure 5.8), we determined the point at which the mood state (i.e., symptom sum score) was closest to the midpoint $J/2$ (i.e., half of symptoms activated). Subsequently, the hysteresis effect was computed by subtracting the point at which the increasing stress line was closest to the midpoint from the point at which the decreasing stress line was closest to that same midpoint. A value of 0 thus indicates no hysteresis (point at which symptom sum score is closest to the midpoint is equal for both increasing as well as decreasing stress) while increasing positive values H indicate a larger hysteresis effect: the larger H, the more stress reduction is needed to reduce the sum score to roughly half the symptoms when stress is decreasing; compared to the amount of stress that results in the activation of roughly half of the symptoms when stress is increasing. For every 20 simulations with the same parameter values, we computed the average hysteresis effect on which subsequent analyses were based.

Figure 5.9 shows the average hysteresis effect for the five parameters whose values we varied systematically. First, with linear models, we estimated the impact of each parameter on the hysteresis effect. Most importantly, the connectivity of the network significantly impacted the amount of hysteresis (as we found in our main simulation; estimate: 3.75, $t = 11.02, p < .001$): the more strongly connected the network, the larger the hysteresis effect. Additionally, except for the b parameter, all other parameters influenced the hysteresis effect significantly: continuous weights result in a more pronounced hysteresis effect compared to discrete weights (estimate: 1.93, $t = 3.47, p < .001$); the more nodes in the network, the stronger the hysteresis effect (estimate: 0.70, $t = 9.94, p < .001$); and the higher the mean of a, the stronger the hysteresis effect (estimate: 4.10, $t = 6.02, p < .001$). Finally, two interaction effects (computed with all variables being centered around their respective means) were significant as well: the effect of connectivity on the hysteresis effect was even stronger for an increasing (1) number of nodes in the network (estimate: 0.40, $t = 6.57, p < .001$) and (2) mean of the a parameter (estimate: 2.59, $t = 3.09, p < .001$). As such, we conclude that the hysteresis effect is robust in that increasing connectivity of a network results in more hysteresis; and that this effect is more pronounced for networks with more nodes and a probability function with a higher average a parameter.

Taxa versus continua

The question whether psychopathological conditions are instances of taxa (i.e., distinct categories, natural kinds) or continua (i.e., a continuous dimension with psychopathology located at an extreme of that continuum) is an old one and one without a definitive answer (e.g., Flett, Vredenburg, & Krames, 1997; R. E. Kendell, 1975; R. Kendell & Jablensky, 2003; Haslam, Holland, & Kuppens, 2012; Meehl, 1992; Waller & Meehl, 1998; Widiger & Samuel, 2005). There is evidence for both views although the last few years, more evidence seems to point into the direction of psychopathological conditions being continua (e.g., Haslam et al., 2012), which poses a challenge for diagnostic schemes, such as the DSM that offer categorical perspectives on psychiatric diagnosis. Pertaining to major depression specifically, the results are inconclusive: for example, Baldwin et al. (Baldwin

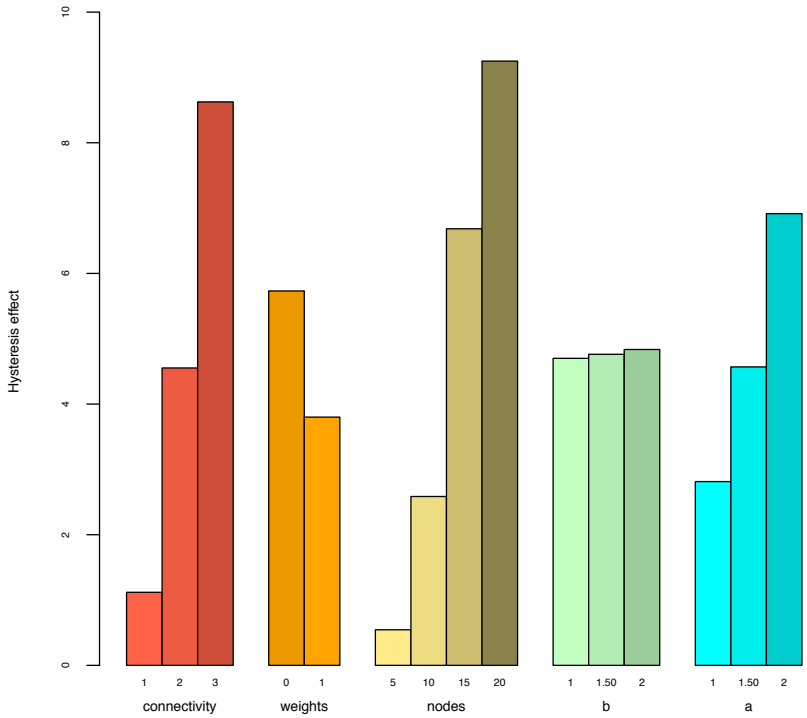

Figure 5.9: The average hysteresis effect for different values of five parameters of the MD simulation model: (1) *weights*: weights of the connections between nodes, either continuous (0) or discrete (1); (2) *connectivity*: parameter with which the adjacency matrix is multiplied, either 1, 2 or 3; (3) *nodes*: number of nodes, either 5, 10, 15 or 20; (4) b: b parameter in the probability function randomly sampled from a normal distribution with mean 1 ($\sigma^2 = 0.25$), 1.50 ($\sigma^2 = 0.50$) or 2 ($\sigma^2 = 0.75$); (5) a: a parameter in the probability function randomly sampled from a normal distribution with mean 1 ($\sigma^2 = 0.25$), 1.50 ($\sigma^2 = 0.50$) or 2 ($\sigma^2 = 0.75$).

& Shean, 2006) report evidence, based on taxometric procedures, that total scores on the Center for Epidemiological Studies Depression scale (CES-D) were best represented as a continuum; while Ruscio et al. (Ruscio, Zimmerman, McGlinchey, Chelminski, & Young, 2007) report evidence that total scores on a semi-structured clinical interview for assessing major depression were best represented as a taxon. The results of our simulation offer an alternative stance in the taxon-continuum debate, namely that MD is *both* a taxon and a continuum, depending on the connectivity of the network. More specifically, as one can infer from Figure 5.8, networks with weak connectivity appear to behave as continua: transitions from less to more MD symptoms and vice versa take place in a smooth, continuous fashion. That is, the network model for MD hypothesizes that in people resilient to develop episodes of MD, progressing from not having symptoms to many symptoms, and back, under the influence of stressors, is a continuous smooth process. On the other hand, networks with strong connectivity appear to behave as taxa: under the influence of stress, transitions from less to more MD symptoms and vice versa take place in a discontinuous with a 'forbidden' zone of rarity (between 1 and 5 symptoms) that does not seem to function as a stable state. This hypothesis is consistent with empirical findings that major depression appears to be a taxon in patient samples

while a continuum in studies with community samples (which often contain many people that are resilient for developing depression; e.g., Baldwin & Shean, 2006; Ruscio et al., 2007; Ruscio, 2009; Slade & Andrews, 2005; Slade, 2007). Additionally, our hypothesis sits well with clinical observations that for patients prone to developing episodes of MD, their mood states indeed seem to fall in two distinct zones with a large gap between the two.

Diathesis-stress theory of major depression

The generic diathesis-stress model (Abramson, Metalsky, & Alloy, 1989; Bebbington, 1987; Beck, 1987; McGuffin, Katz, & Bebbington, 1988; Robins & Block, 1988) attempts to answer questions such as why some people develop MD after experiencing stressful life events while others do not. The model does so by positing that developing disorders such as MD is the result of an interaction between a certain *diathesis* (i.e., vulnerability) and a range of possible stressors. According to this model, the experience of a certain stressful life event can activate the diathesis, thereby "...transforming the potential of predisposition into the presence of psychopathology" (Monroe & Simons, 1991). Thus, the theory posits that some people are just more vulnerable than others and when put under stress, these people have a high risk of developing MD and experiencing relapse. Exposed to the same stress, those without this diathesis are at quite low risk.

A substantial problem with existing formulation of the diathesis-stress model is that despite many efforts to define diathesis, there is no universally accepted and proven hypothesis about what diathesis is: what do we mean when we say that a particular person is vulnerable? Does diathesis/vulnerability refer to some pathology that is predominantly (1) biological such as having lots of risk alleles, (2) developmental like having experienced sexual abuse as a child or (3) psychological such as having a relatively high level of neuroticism (Caspi et al., 2003; Ensel & Lin, 1996; Harris et al., 2000; Kessler & Magee, 1993)? Or, alternatively, is vulnerability a particular combination of these three pathologies, for example being sexually abused as a child which causes relatively high levels of neuroticism later in life? Additionally, there is no consensus about how diathesis and stress interact in bringing about an episode of MD (Belsky & Pluess, 2009; C. Hammen, 2005; Monroe & Simons, 1991): do stressful life events activate the diathesis (e.g., loss of a loved one results in the expression of risk alleles) or does diathesis act as a moderator (e.g., loss of a loved one results in the development of more MD symptoms in someone with lots of risk alleles than in someone without these risk alleles)?

Our network model of MD *is* a diathesis-stress model in that it both explicitly models vulnerability—which is, in our model, the overall strength of the connections in a given network—as well as the interaction between this vulnerability and the influence of stressful life events: stress augments the probability of symptoms to become 'active' resulting in one or a few symptoms becoming actually active; and when the connections in the MD network are generally strong, then these activated symptoms will result in the activation of other symptoms as well, eventually culminating in a full-blown episode of MD. As such, our model provides an alternative, concrete and testable interpretation of diathesis, stress, and the interaction between them; and our preliminary results are encouraging for this particular interpretation of the diathesis-stress model. Additionally, the results of our simulation show that the level of stress necessary to activate symptoms was decreasing with increasing connection strengths. As such, an additional testable hypothesis, for instance via time series modeling, would be that when a person is more vulnerable to develop an episode of MD (i.e., stronger intra-individual connections), less stress is needed to induce the activation of symptoms (possibly culminating in a full-blown episode of MD).

Treatment

The hypothesis, which we confirmed with simulations, that vulnerable networks (i.e., networks with strong connectivity) behave as a cusp catastrophe model potentially offers a new outlook on treatment: more specifically, the different therapeutic approaches for patients with MD can be categorized according to their effects on the MD system. First, a continuous *force* may be applied that reduces the whole set of stress parameter values. The aim would be to reach the lower left threshold that shifts the mood system back towards the healthy attractor. Many antidepressants may work through this mechanism. The metabolism or reuptake of monoamines (e.g., serotonin and norepinephrine) is blocked immediately after drug administration, while different antidepressants were shown to positively affect emotional processing in healthy subject already within one week (A. Frazer & Benmansour, 2002; Harmer, Shelley, Cowen, & Goodwin, 2004; Harmer, 2010). Yet, in many patients it takes more than 6 or 8 weeks to experience the full benefits of treatment with antidepressants (Quitkin, McGrath, Stewart, Taylor, & Klein, 1996; Trivedi et al., 2006). The hysteresis effect may help to explain the delay in antidepressant drug action. At critical thresholds, tiny disturbances may not only cause a large affective shift (i.e., a critical transition), but that shift is accompanied by a high degree of irreversibility (i.e., inertia). That is, the system can shift towards MD under the influence of a mild stressor, but if that same amount of stress is subsequently removed from the system, the system does not directly return to a healthy state. Second, a strong *perturbation* may be applied that 'kicks' the mood state of the MD system. In other large-scale complex systems with alternative attractors and tipping points, perturbations may tip the system into an alternative basin of attraction (Scheffer et al., 2009; van Nes & Scheffer, 2007). A perturbation increases the chance of arriving at another (more desirable) attractor. Sleep deprivation, the *N*-methyl-*D*-aspartate (NMDA) antagonist ketamine, and electroconvulsive therapy may destabilize a person's basin of attraction, which may induce rapid (but often transitory) antidepressant effects (J. C. Gillin, Buchsbaum, Wu, Clark, & Bunney, 2001; Zarate et al., 2006). A third option from a theoretical standpoint would be to loosen or split the connections between the nodes of the MD network. This would transform the whole shape of the state space landscape, removing bifurcations and the hysteresis effect. Depressive symptoms would recede when the adverse influence of one activated node of the node would no longer transmit its effects to other nodes, and a more resilient state would be achieved. Cognitive behavioral therapy contains techniques that can help to desynchronize and loosen the connections between MD symptoms (see also Cramer et al., 2010): techniques that help a patient, when experiencing, say, depressed mood, to not easily let that depressed mood cause suicidal thoughts, for example by challenging the patient's assumptions about the abnormality of suffering from depressed mood from time to time. Finally, combining these options might be most promising: loosening the connections between MD symptoms first, followed by reducing stress, may allow for a continuous and smooth path to euthymia. However, we are still at the first stages of exploring the use of the cusp catastrophe model in MD and therefore, these treatment options are merely theoretical at this point.

Discussion

We have shown that a model in which MD is characterized as a network of causally connected symptoms has the potential to explain what makes certain people vulnerable to develop an episode of MD: the stronger the connections between the symptoms of someone's individual MD network, the easier it is for a full-blown episode of MD

to develop. Additionally, we showed that in the weakly connected networks (i.e., the hypothesized resilient networks/people) spontaneous recovery occurred, a well-known clinical phenomenon. The network model of MD also has the potential to explain why some people develop an episode of MD after (mild) stress while others do not; and why the shifts from depressed to healthy states and vice versa generally follow a non-linear pattern. More specifically, we have formulated a novel definition of diathesis in terms of the strength of connections between MD symptoms; as well as a novel hypothesis about the interaction between diathesis and stress: stressful life events influence individual symptoms directly and the diathesis then determines whether a cascade of symptom development emerges that can culminate in a depressed state. With simulated data, we have shown that our formulation of the diathesis-stress model works: resilient networks (i.e., with weakly connected symptoms) could handle significant amounts of stress without falling into a depressed state; while the vulnerable networks (i.e,. with strongly connected symptoms) behaved like a cusp catastrophe: at tipping points, only slight amounts of stress sufficed to tip these networks into a depressed state.

There are some extensions of the model that might serve to test additional theories about the pathogenesis and maintenance of MD. For example, there is evidence for a reverse relation between stressful life events and MD in which the presence of depressive symptoms predisposes a person towards experiencing certain stressful life events (Maciejewski, Prigerson, & Mazure, 2000): e.g., experiencing fatigue and loss of interest resulting in the loss of employment. Our model could thus be extended by allowing symptom development to trigger the development of stress. Additionally, *kindling*—the phenomenon that stressful life events play the greatest role in the first onset rather than in subsequent episodes of MD—might be incorporated in the model by making the connections between MD symptoms stronger after every MD episode: that is, each consecutive episode makes the network more vulnerable. This extension of the network model would resonate with evidence suggesting that kindling does not necessarily mean that after the first onset, subsequent episodes come out of the blue; rather, such episodes are elicited by less and less severe life events (Monroe et al., 2006). As Kraepelin (1921) noted about one of his patients: she became depressed "after the death first of her husband, next of her dog, and then of her dove" (pp. 179). A final future extension of our model could be the incorporation of the fact that not all symptom dynamics take place within the same time scale. For example, it stands to reason that mood is a variable that can fluctuate within a time scale of hours; but that the relation between insomnia and fatigue unfolds over days (in most people, insomnia one day will not immediately cause fatigue the next day). And the relation between depressed mood and thoughts of suicide is likely an even slower dynamical process. These different time scales are currently not implemented in the model. Additionally, it might also be another source of intra-individual differences: it is possible that vulnerable people differ from healthy people in that the dynamics of the former group are on a faster time scale than the dynamics of the latter group. For example, it is possible that, ceteris paribus, in vulnerable people, depressed mood causes thoughts of suicide faster than in resilient people.

In this chapter, we used simulated data but naturally, testing the model with empirical data is needed in order to draw definitive conclusions. What kind of data would we need for such testing? We cannot stress enough that time-intensive data is key in testing many, if not all, assumptions of network(-like) models. More specifically, one would need intra-individual data in which individuals are followed for a long period of time and are asked about life events, minor daily hassles and symptoms at multiple time points per day. Experience Sampling with data collection through electronic diaries and smart phones would be a particularly suitable method for collecting such data (Myin-Germeys et al., 2009). With such data, it becomes possible to estimate network parameters for

individual people and to test whether intra-individual networks with strong connections are indeed vulnerable to the impact of relatively mild stressors. Also, one would expect that resilient people, when suffering from some symptoms due to a stressful life event, have a higher probability of recovering spontaneously than vulnerable people. Finally, one could also further test the hypothesis that in vulnerable people, MD is a taxon while it is a continuum in resilient people.

If an extended model based on empirical data would confirm that vulnerable networks (i.e., networks with strong connections) behave according to the dynamics of a cusp catastrophe, what would the implications be? For one, a cusp catastrophe implies the existence of tipping points. Finding these tipping points for individuals' networks could then prove beneficial for two reasons. First, knowing that someone's MD system is close to tipping from a healthy to a depressed state would allow for precisely timed therapeutic interventions that might prevent such a catastrophic shift. Second, on the other hand, knowing that someone's MD system is close to tipping from depressed to a healthy state would offer the opportunity of giving the system a large kick (e.g., sleep deprivation) *at exactly the right time* so that the system, like the ball in the bowl, is kicked out of the depressed attractor and ends up in the healthy attractor. Thus, the tipping points in the cusp catastrophe model might help in predicting when prevention and intervention have the highest probability of success.

But how does one find these tipping points? Recent findings suggest that all catastrophic systems, from financial systems to the climate, display *early warning signals* that a system is approaching a tipping point (Carpenter & Brock, 2006; Dakos et al., 2008; Fort, Mazzeo, Scheffer, & van Nes, 2010; van Nes & Scheffer, 2007; Scheffer et al., 2009). One such early warning signal is called *critical slowing down*: right before a tipping point, the system is getting slower in recovering from small perturbations. Pertaining to MD, for instance, one might see that someone has more difficulty than usual to recuperate from a minor daily hassle like an unpleasant day at work. Numerically, this slowing down can be traced by inspecting *autocorrelations*: the correlation between scores of the same variable at multiple time points (e.g. the correlation between 60 measurements of depressed mood). Such autocorrelations go up when the system slows down: slowing down means that at each time point, the system much resembles the system as it was at the previous time point, meaning that the autocorrelation is relatively high. With time-intensive intra-individual data it will become possible to inspect autocorrelations and other potential signals that someone is in critical danger of developing an episode of MD or that a healthy state is within reach.

Chapter 6

Dimensions of normal personality as networks in search of equilibrium

Abstract

In one currently dominant view on personality, personality dimensions (e.g., extraversion) are causes of human behavior, and personality inventory items (e.g., "I like to go to parties" and "I like parties") are measurements of these dimensions. In this view, responses to extraversion items correlate because they measure the same latent dimension. In this chapter, we challenge this way of thinking and offer an alternative perspective on personality as a system of connected affective, cognitive and behavioral components. We hypothesize that these components are not associated because they measure the same underlying dimension; they do so because they depend on one another directly for causal, homeostatic or logical reasons (e.g., if one does not like people it is harder to enjoy parties). From this *network perspective*, personality dimensions emerge out of the connectivity structure that exists between the various components of personality. After outlining the network theory, we illustrate how it applies to personality research in four domains: 1) the overall organization of personality components, 2) the distinction between state and trait, 3) the genetic architecture of personality; and 4) the relation between personality and psychopathology.

Adapted from: Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., Kendler, K. S., & Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality, 26*, 414-431.

People differ widely in how they navigate through the landscape of life: some people feel comfortable around others and like to go to parties, whereas others do not. Some worry much and often have trouble sleeping, whereas others rarely experience such problems. Two main challenges in personality psychology are (1) to provide a plausible account of how the coherent 'organization' of such behaviors arises within an individual and (2) to describe and explain the structure of 'individual differences' in personality (Caprara & Cervone, 2000). Modern personality psychology has mainly focused on the latter task. Starting with pioneering work of, among others, Thurstone (1934), the dominant doctrine in current personality theory has come to define individual differences in the structure of personality in terms of a number of unobserved trait 'dimensions' (e.g., neuroticism and extraversion; Berrios, 1993; Digman, 1990; Goldberg, 1993).

In most interpretations of this concept, consistent differences between people in the behavior they display are thought to result from underlying differences in these personality dimensions. The interpretation of the term 'underlying' is typically borrowed from measurement models in psychometrics, which invoke latent variables —variables that are indirectly measured through a number of noisy indicator variables (i.e., personality inventory items; Borsboom, 2008b). In line with this mode of thinking, the items of a personality inventory are usually considered to be 'trait measurements', for example, the item "I like to go to parties" is considered a measurement of the dimension/trait extraversion. Analogous to temperature, which causes mercury to rise and fall in a mercury thermometer, personality dimensions are presumed to cause responses to personality questionnaire items. For example, higher levels of extraversion cause people to make friends more easily and to feel good in the company of others, and these properties are queries in typical questionnaire items. Thus, for example, Alice is not only more extraverted than Bob in the sense that her responses can be 'described' by a higher position on an abstract personality dimension (e.g., extraversion); her higher level of extraversion is what 'causes' her to like parties better than Bob does. In this way, personality dimensions are interpreted as causes of human behavior. Perhaps the most outright commitment to this point of view is expressed in McCrae and Costa (2008, p. 288) who claim that "E[xtraversion] causes party-going".

In this chapter, we challenge this approach to personality. In addition, we offer an alternative perspective. We propose that personality is a system of inter-connected affective, cognitive, and behavioral 'components'. More specifically, we propose that every feeling, thought or act is a potential component of personality if it is associated with a unique 'causal system': the pattern of causes and effects that the component exhibits in relation to other components. The component must thus be unique in the sense that its causal system differs from that of other (potential) components. This means that a personality component is, to a certain degree, causally autonomous and, as such, not 'exchangeable' with other components. Thus, liking parties is a personality component because it has unique causes and effects on other components (e.g., being interested in meeting new people and not feeling insecure about making a good first impression) that differ from the causes and effects of other components (e.g., starting conversations easily, also an extraversion item, does not necessarily imply that one is interested in meeting new people). To the contrary, making to-do lists that are followed point by point and sorting clothes by color may not be separate components at the level of personality (i.e., their causes and effects on other components will likely be similar) but two ways of assessing one component, namely liking order. Barring such exceptions, personality components

are typically assessed through single items in personality inventories[1]. This is because two items that assess precisely the same component will be very highly correlated, which tends to cause problems in typical psychometric analyses (i.e., this will show up as correlated errors).

We hypothesize that such components cannot change independently of one another and, therefore, form a network of mutual dependencies that may alternatively have causal, homeostatic or logical sources. Directional dependencies (typically associated with 'causality') will form if one component influences the other but not the other way around: for example, if one cannot plan ahead, it is difficult to meet obligations at work. Bidirectional dependencies will form if two components influence one another (and, as such, create a feedback loop): for example, after a sleepless night worrying, one may feel stressed out and tired the next day; as a result of which one may not sleep the following night either because of worries about yet another sleepless night. An important special case of feedback involves negative feedback loops that serve to maintain 'homeostasis': for instance, after a few sleepless nights, one will ordinarily get so tired that one will start sleeping again (incidentally, if this does not happen, problems are likely to spread to other components, e.g., not being able to concentrate and foul mood; Cramer et al., 2010). Finally, semantically 'logical' dependencies will form if two components assess the same or a narrower/broader version of a personality characteristic (which may ultimately but not necessarily result in these two components being merged into one component): for example, liking a clean house and liking a clean desk. We postulate that the resulting pattern of connectivity among such components provides a fruitful avenue into personality research. Also, the dependencies between these components result in a typical network architecture (e.g., being interested in other people and spending time with them are mutually dependent while planning ahead and liking people are not) that can serve as a sufficient explanation of the correlational structures typically observed in personality research (e.g., trouble falling asleep and feeling jittery are more strongly correlated than feeling jittery and liking people).

This opens the perspective of a personality theory that is holistic (i.e., that is about the 'organization' of behavior: network architecture) and that addresses personality at the level of the individual but is nevertheless systematically formalizable through network models. Importantly, this view does not regard personality dimensions as causes of behavior. We will instead argue that personality dimensions emerge out of the connectivity structure that exists between its components, such that certain components cluster together more than others, with the known personality dimensions as a result.

The structure of this chapter is as follows. First, we examine affective, cognitive and behavioral components of personality and argue that a network perspective naturally accommodates mutual (in)dependencies among them (see also Cramer et al., 2010;

[1] We acknowledge that personality inventories tend to measure self-concept, a person's view of one's own personality that might, to some extent, deviate from one's actual, objective personality. Because the network perspective is undecided concerning whether or not personality networks should be solely based on objective personality, or on both objective personality and self-concept, using personality items as a starting point for defining personality components is sensible, also given the lack of viable alternatives. Future experimental research with a focus on elucidating whether thoughts/feelings/acts, and their mental representations, have unique effects on other personality components might prove beneficial in refining personality networks in terms of what components they should contain: objective ones or also their mental representations. However, we note that current latent trait models often do equate self-concept and objective personality: the personality literature shows an abundance of statement concerning traits (e.g., women are more neurotic compared with men, suggesting an objective difference); the evidence for which is often based on personality inventory items (a more appropriate statement would then be as follows: women's self-concept of their personalities tends to include more neurotic features compared with men).

Schmittmann et al., 2013; Read et al., 2010 for similar perspectives). We describe the consequences of the network perspective for prominent topics in personality psychology in the subsequent sections of this chapter, which deal with the state-trait distinction, the relation of personality to psychopathology and the genetic basis of personality. Throughout the chapter, we relate the network perspective to currently dominant trait theories.

Humans as dynamical systems

Few psychologists would challenge the conclusion that human beings are complexly organized. Even the simplest behavioral act (e.g., starting a conversation with a stranger while waiting for the bus) reinforces cognitive schemas (e.g., it provides evidence for the hypothesis that one is capable of starting such a conversation) and affective conditions (if the small talk is successful, this most likely generates a feeling of satisfaction). Because these cognitive and affective components are associated with a class of behaviors in a given situation, they almost certainly serve to sustain the ability and willingness to execute these behaviors when a similar situation is encountered in the future (Mischel & Shoda, 1995). That is, one who has successfully engaged in small talk and enjoyed it is likely to engage in small talk again.

Thus, even this extremely simple example suggests the presence of feedback loops among the components of the personality system, in which behavior is not just an outcome variable in need of explanation but itself may serve as input to the system (i.e., the behavior was successful so probably will be executed again under similar circumstances in the future). There most likely are many such feedback mechanisms operating at different time scales, giving rise to a dauntingly complex picture. Thus, Skinner (1987) was definitely on target when he said that human behavior is 'possibly the most difficult subject ever submitted to scientific analysis'. In fact, in view of the stunning complexity of the system, it should be considered remarkable that stable behavioral patterns exist at all.

But they do. For some reason, human systems tend to settle in relatively fixed areas of the enormous behavioral space at their disposal, where they are in relative 'equilibrium' with themselves and their environments. By equilibrium, we mean a stable state (e.g., Joan is interested in other people and sympathizes with their feelings; as a result of which, she has a job as a social worker) that is not left upon a small disturbance (e.g., one of Joan's clients steals some money from her; after which, she is naturally disappointed in the culprit, but she is still interested in people and their feelings, and she continues her job as a social worker). This definition of equilibrium is analogous to 'attractors' in the complex systems literature (e.g., Teschl, 2008).

The idea that human beings strive to survive and reproduce by actively interacting with their environments is an old one and can be traced back to Darwin (1871). In psychology, several scholars have argued for a theory in which human beings are open systems that are constantly searching for equilibrium or a state of homeostasis. Such equilibriums have been argued to exist with respect to components internal to the human system (e.g., in Freudian psychology, id and superego) and with respect to the relation between the human system and its environment (Allport, 1960; Stagner, 1951; Tryon, 1935).

Such states of homeostasis, which we designate to be 'behavioral equilibriums', can be achieved and maintained in several ways. For instance, people can and will (consciously or not) seek out environments that match their behavioral repertoire (e.g., Heady & Wearing, 1989; Kendler & Baker, 2007; Kendler, Gardner, & Prescott, 2003). For instance, Alice who loves to go to parties, actively seeks environments that provide many opportunities to party and to meet people who can invite her to parties. Thus, organism-

environment feedback loops are important sources of stability because they can serve to sustain behavioral patterns. As a consequence of such feedback-driven selection of environments, however, other behavioral states can also become more difficult to access. This is because they would require different types of environments. For example, Alice cannot both love parties and dislike being around people at the same time. For the former preference to thrive, a socially busy environment with many parties is required, whereas the latter preference would require a more tranquil environment featuring only a limited number of people. Thus, active selection of environments has two important consequences. First, it allows people to settle in an 'typical' pattern of behavior (a behavioral equilibrium, analogous to an 'attractor' in complex systems theory) through organism-environment feedback loops. Second, it creates negative dependencies between behaviors that require different environments because people's behavioral options are not inexhaustible: every behavioral act comes at the expense of not performing another one and, as such, closes the futures that could have been if another act had been chosen.

It is further characteristic of the behavioral patterns typically studied under the rubric of personality that they can be shaped and maintained in a variety of ways. Thus, people can respond in their own idiosyncratic ways to the situations in which they find themselves. For example, Jane does not like being around people she does not know very well, so when an acquaintance throws a party, she will attend but she will not mingle much and go home as early as is politely possible. However, at a family reunion, she enjoys the company of her close relatives and stays late to catch up with them. These idiosyncratic patterns of situationally dependent responses have been addressed in the cognitive-affective personality system, which we consider to be naturally compatible with a network perspective (CAPS; Mischel & Shoda, 1995, 1998). According to the CAPS model, personality depends not only on the person but also on the environment, that is, one's idiosyncratic way of behaving is stable within environments but variable across environments.

As a result, each person defines a somewhat idiosyncratic equilibrium with his or her environment that is likely to be organized around some properties that play key roles in the individual's cognitive and affective economy (i.e., that are important to the person; Cervone, 2005). Because of the connectivity structure of the human-environment system, these properties cannot vary entirely in isolation: one is unlikely to enjoy parties if one is nervous around others, and one cannot be nervous around others if company makes one feel comfortable. Similarly, in the realm of conscientiousness, one cannot be completely successful at finishing tasks in time if one cannot plan ahead, and for finishing tasks, its generally helps if a person enjoys working hard. Some of these properties are connected, in the sense that they are mutually dependent on one another. Other properties are unconnected or very weakly connected (i.e., relatively mutually independent). For instance, one can like working hard without being able to make friends easily. Thus, these dependencies between personality components define the structure of the network that characterizes a person: the personality architecture (Cervone, 2005).

Now, suppose that one settles into a behavioral equilibrium with respect to one property. Say, a person likes to be around people and as a child seeks the company of others systematically (for a similar point of view, see Caspi, Bem, & Elder, 1989; Caspi, Elder, & Bem, 1987, 1988). As a result, one's social skills are developed and improve over time, which makes it easier to be around others, until at some point an equilibrium is reached. This means that the situation has become relatively stable: one likes to be around others, and one has succeeded in findings a way to realize that state (e.g., a job in a social environment), barring situations where one is temporarily and involuntary 'kicked out' of equilibrium (e.g., being ill and therefore unable to leave the house for some time). Then, the evolution of this property (i.e., enjoying the company of others)

will cause other properties, such as social skills, to co-evolve into a related equilibrium: it is quite hard to like to be around people and actively seek out environments that match this preference without at the same time developing social skills. Another person may reach the same equilibrium but approach from the other direction; for some reason, the person becomes highly skilled in social interactions and comes to like the company of people as a result. Thus, groups of properties will move synchronously, like a flock of birds or a swarm of bees, simply because the organization of the human system and its environments require it.

This idea stands in stark contrast to the idea that behavior is caused by a small set of latent personality dimensions/traits. In terms of the flock of birds analogy: in the situation as mentioned earlier, one bird in the flock flies in a particular direction because its neighboring birds do so; in a latent trait scenario, all birds in the flock fly in a particular direction because of the instructions of an invisible (i.e., latent) bird. That is, in the standard model, personality dimensions/traits function as 'common causes' of a set of item responses (Borsboom, 2008a; Edwards & Bagozzi, 2000; Schmittmann et al., 2013). In psychometric terms, one of the most important features of a latent trait model that signals this assumption is 'local independence' (Holland & Rosenbaum, 1986; Lord, 1953; McDonald, 1981). Local independence means that, conditional on any given position on the latent variable, the observed item responses are statistically independent. Essentially, this means that the associations between items are spurious in the sense that they arise 'solely' from the items' common dependence on the latent variable. This is structurally analogous to the textbook example of the correlation between the number of storks and the number of newborns across Macedonian villages: villages that have more storks also have more newborns. This association is spurious because the correlation between storks and newborns arises solely from both variables' dependence on village size: larger villages have more chimneys, which attracts storks, and more people, who produce babies.

A latent variable model works in the same way. It relies on the assumption that dependencies among the cognitive, affective and behavioral components of personality (i.e., the individual birds in the flock, for example neuroticism items) arise 'solely' because all components depend on the same underlying trait (i.e., the invisible bird, for example neuroticism). Figure 6.1 shows an application of this model to the Big Five dimensions as measured with the NEO-PI in 500 first year psychology students at the University of Amsterdam (Dolan, Oort, Stoel, & Wicherts, 2009). Reliance on the assumption of local independence is evident by the absence of any direct connections between items. As such, local independence explicitly prohibits causal relations between the components of personality as represented by the items. The model with five latent traits influencing only their respective items, as depicted in Figure 6.2, does not fit the data ($df = 28430$, $\chi^2 = 60839$, $p < 0.001$), which is mainly due to violations of simple structure: particularly, the correlations between items that belong to different personality dimensions are too high to be accounted for by the model. How can one address this problem? One way, the standard way in personality psychology (e.g., Savla, Davey, Costa, & Whitfield, 2007) is by tweaking the model 'on the basis of the data' so that the basis latent variable hypothesis is preserved (e.g., by allowing cross-loadings, exploratory factor analysis with procrustes rotation; see also Borsboom, 2006 for an elaborate critique). Another way would be to make the simple structure more complex, for example, by introducing first-order and second-order latent variables (not to detract from the main aim of the chapter, we have not included fitting such more complex models). Another way, the central tenet of this chapter, is to consider the misfit of the untweaked model an indication that the latent variable hypothesis fails as an explanation of the emergence of normal personality dimensions, and to move on towards alternative models.

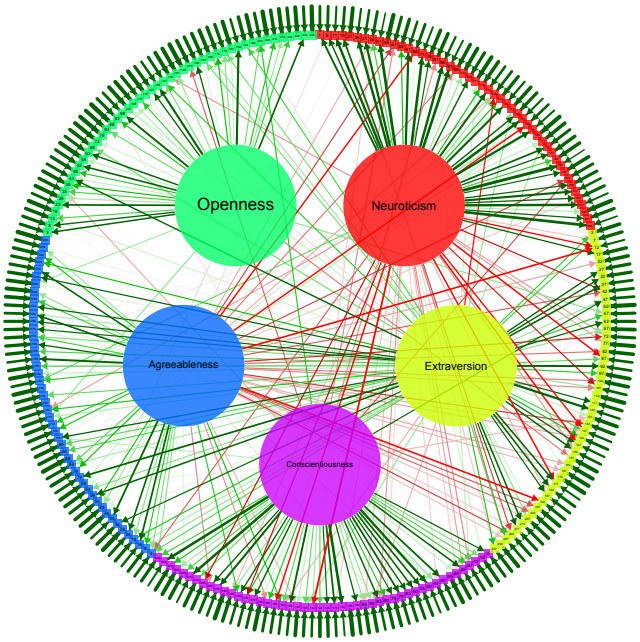

Figure 6.1: The five-factor model for the NEO-PI items. Covariation between items is explained by the hypothesis that five latent variables (big circles in the middle) act on distinct sets of items (boxes). Positive parameters in the model are green; negative parameters are red. The arrows between circles and boxes represent factor loadings, arrows between circles are correlations between factors and arrows pointing into the boxes represent residual variance.

That is, because of the local independence assumption, the very idea of cognitive, affective and behavioral components (i.e., items) that are directly connected to one another for causal or homeostatic reasons (or, for that matter, because of logical ones) is irreconcilable with the dominant latent trait perspective on personality dimensions and their items. If we take the connections between the components of personality to be real, that is, non-spurious, then a viable alternative approach is to describe them as a network. The crucial aspect of such a network is its organization: the way in which functional components of human personality are linked to one another. In turn, this organization depends critically on the equilibriums of the human system and its environments: certain behaviors correlate or coincide, whereas others do not because they are compatible or incompatible, respectively, with respect to specific equilibriums.

From this point of view, neuroticism items are tightly connected not because they are caused by the same latent trait (neuroticism) but because they arise in similar equilibriums: for example, someone who feels threatened easily will likely also suffer from nerves, feel lonely and worry too long after an embarrassing experience. Items related to the free exploration of environments (e.g., being open to new people) will unlikely co-evolve within threat-related equilibriums and hence will not be tightly connected to neuroticism items. This is not to say it is no longer valid to speak of 'neuroticism' or

'openness' as personality dimensions/traits: it certainly is, but under the assumption of a network perspective, these terms do not indicate latent causes of behavior but groups of tightly inter-connected personality components. Thus, we can still use a term such as 'neuroticism' to refer to a phenomenon that emerges as a result of the biological, psychological and environmental forces that knit some behaviors closely together. However, we speak of such a phenomenon just like we speak of a flock of birds. We know that a flock emerges out of the synchronized behavior of the birds it contains and would not venture hypothesize that it existed independently of that behavior, let alone was caused by it (Schmittmann et al., 2013).

Naturally, we are not the first to raise questions about the incompatibility of current trait models with dynamic interactions between personality components and the environment. Similar ideas have been manifested in the writings of personality theorists, almost since the inception of the discipline; recent theorists such as Mischel and Shoda (1995) and Cervone (2005), as well as Read et al. (2010), have argued along very similar lines. However, the methodology to study complex networks has been developed to maturity only relatively recently (R. Albert & Barábasi, 1999; Newman, 2006; Watts & Strogatz, 1998). As a result, we are now able to use such techniques to visualize and analyze large-scale networks in ways that have not been possible before. The remainder of this chapter aims to give first passes at applying these ideas systematically in the study of personality. We focus on four illustrations regarding (1) the overall organization of personality components, (2) the distinction between state and trait, (3) the genetic architecture of personality and (4) the relation between personality and psychopathology.

A network of personality components

Mapping the structure of personality onto a network is a daunting task. Fortunately, we have a reasonable starting point in the form of common personality questionnaires that query respondents for their status with respect to exactly the type of components that would be likely candidates to make up a personality network structure. The correlations between components will tend to be higher when the connectivity in the human system is stronger. Thus, by studying correlations and by representing them in a network structure, one may obtain a first glance at the visualization of the global (i.e., average) structure of personality components. We have developed an R-package for network analysis (qgraph: Epskamp et al., 2012) that is capable of constructing such visualizations directly from the data. In essence, the routines in this package treat a correlation matrix as a so-called weighted network, that is, it treats the items as components and their correlations as the strength of the connections among these components. The result of applying this routine to the items of the NEO-PI-R is presented in Figure 6.2 (for the large central graph, same sample as used for Figure 6.1; for the small graph in the top right, simulated data)[2].

A graph like that in Figure 6.2 offers a powerful visualization that can be used to reveal patterns and structures that would be very difficult to spot by using traditional methodology (note that Figure 6.2 represents the complex structure of no less than 240 x 240 = 57600 correlations with little data reduction). For instance, looking at Figure 6.2, there are a few things that catch the eye immediately. First, the network is very

[2] Please note that for this graph, and the other networks that are presented in this chapter, the positions of the nodes in the graph are not identified. That is, by using a force-embedded algorithm, the graphs are a two-dimensional representation of networks that are multi-dimensional. In this representation, the position of a node is defined relative to other nodes in the network. The resulting distance in two dimensions between two nodes does not represent the correlation but, rather, is an approximation of the distances in the multi-dimensional network.

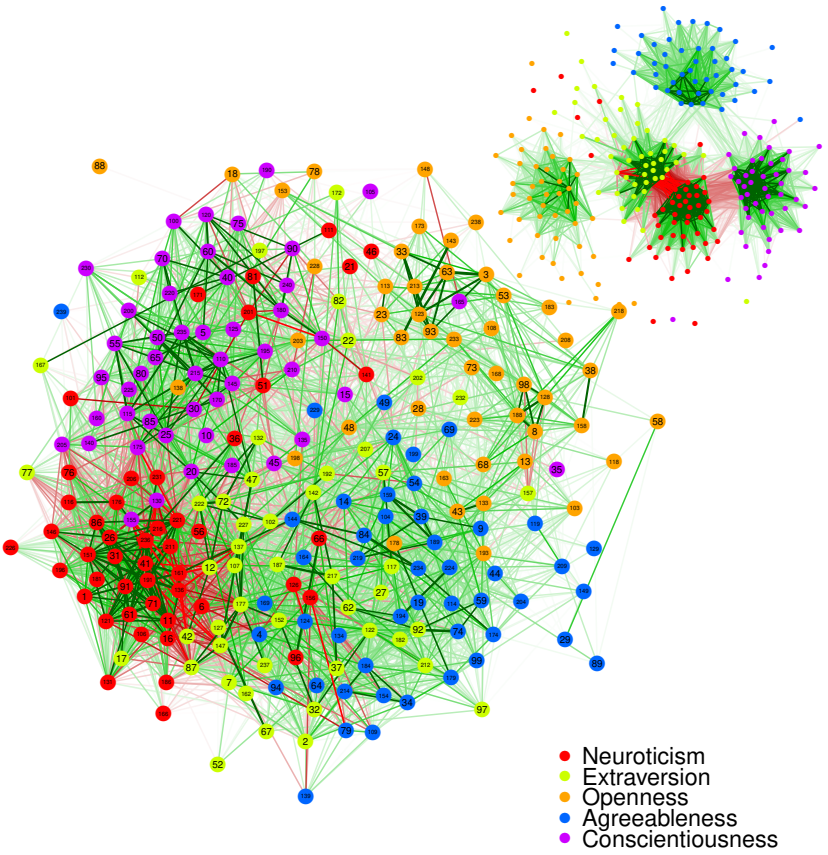

Figure 6.2: A network representation of 240 NEO-PI items based on data (large central graph) and based on expected correlations if a (fitted) five-factor model were true (i.e., simulated data, small graph top right). Each item is represented as a node, and the numbers in the nodes refer to the item numbers in the Dutch version of the NEO-PI. Nodes are connected by green (red) lines if they are positively (negatively) correlated. The thicker the line, the higher is the correlation. The spring-based algorithm used to generate the graph places strongly correlated nodes closely together and towards the middle of the graph.

densely connected, much more connected than would be expected if a small number of latent variables gave rise to the correlational structure (even if we let these five latent variables correlate, as we did in Figure 6.1). In particular, this visualized pattern of correlations between personality items is not convincingly suggestive of five distinct latent traits. This can also be seen when comparing the empirically constructed graph with the inserted graph at the top right of the figure, which shows the correlations that would be expected if the five-factor model of personality were true (i.e., if the covariation between items could be solely explained by five correlated latent variables that cause the item responses).

In this dataset, the strongest organization arises for neuroticism and conscientiousness items (red and purple nodes/circles in Figure 6.2). Extraversion and agreeableness items (yellow and blue nodes in Figure 6.2) are largely intertwined with one another, meaning that, on average, an extraversion item is not much more strongly correlated with other extraversion items than with agreeableness items (and vice versa; difference in average

correlations is 0.12). This makes sense from a network perspective. For instance, it becomes easier to spend time with others (agreeableness) if one likes to be around others (extraversion), and it is difficult to talk much with people at parties (extraversion) when one is not really interested in others (reversed agreeableness item).

Another interesting aspect of the graph in Figure 6.2 is that some items are more strongly connected to other items (those items are placed towards the middle of the graph: e.g., nodes representing item numbers 15, 48, 49, 135 and 229), whereas others are only weakly connected to other items or not connected at all (those items are placed towards the periphery of the graph: e.g., nodes 88 and 239). That is, some items are more 'central' in the network than others (see also Cramer et al., 2010). Without a network representation, one would be very unlikely even to think about a concept such as centrality in this way, let alone think of ways of computing it.

For example, the item "When I promise something, one can count on me to fulfill that promise" (node 135) is a central item in the Big Five network. This makes sense because the content of that item is closely connected to not only other conscientiousness items—for example, to fulfill a promise, one generally has to be a reliable person (node 45) and have a tendency to finish things one has started (node 145)—but also items of other personality dimensions as well (i.e., thick lines in Figure 6.2): for example, someone who likes people and sympathizes with them is more likely to fulfill a promise (agreeableness, node 126) as well as be someone to whom other turn when decisions have to be made (extraversion, node 132). On the contrary, the item "We can never do too much for the poor and the elderly" (node 89) is a peripheral item: other than a few connections with other agreeableness items—people who care about the poor and the elderly generally feel sympathetic towards people are worse off (node 209)—(not) caring about the poor and the elderly has (very) little to do with how open, extraverted, neurotic and/or conscientiousness one is. Thus, items in the Big Five network differ in terms of their centrality in that network, and given the content of the items, these differences in centrality appear to make theoretical sense.

Importantly, the entire notion of central versus peripheral components in the Big Five network is irreconcilable with a latent trait perspective in personality, which is articulated in a latent variable 'measurement' model: in such a model, save for measurement errors, items that measure the same trait are exchangeable and thus equally central or peripheral (factor loadings are reliability estimates and as such, cannot be measures of centrality as we view the concept). For instance, if the latent variable model in Figure 6.1 were true, then someone's position on the conscientiousness continuum could be determined perfectly from knowing the *expected* value of any of the conscientiousness items (Jöreskog, 1971; Lord & Novick, 1968; Borsboom, 2005). In other words, the model holds that if one knew the expected value of a person, say, on the item "I tend to finish things once started", then *none* of the other items would offer *any* additional information about how conscientious that person is (i.e., that person's position on the latent conscientiousness continuum). In that sense, all items are equally central (or peripheral), just like mercury thermometers are no more 'central to temperature' than digital or other thermometers are.

Does it matter if some components in the Big Five network are more central than others? It does because, first, it hints at which pathways are more likely to result in the emergence of certain personality structures in some people. A person's personality structure can be represented in a network analogous to the one in Figure 6.2 (for such an individual network, connection strength then refers to how strongly two personality components are connected over time in one person), a subject we will return to in the next section. Because Figure 6.2 is based on between-subjects data (and is, as such, an 'aggregation' of the networks of all these individual subjects), it is likely that at least in

some of these subjects, the central components in Figure 6.2 are prominent features in their networks as well. The network model predicts that once such a central component becomes 'active'[3] in someone (i.e., a component changes in terms of its state, for example, not having experienced this before, someone starts to experience fear of disappointing others, a component linked to both agreeableness and extraversion; Mongrain, 1993), then the probability of neighboring components to become active as well rises because of the strong connections of that component with other components in the network (e.g., "I get chores done right away" and "I finish things I have started"). This particular pathway (fear of disappointing others ↔ getting chores done ↔ finishing things) to a personality structure in which multiple conscientiousness items are active is then more likely than a pathway to conscientiousness that includes peripheral components (e.g., "I take voting and other duties as a citizen very seriously", node 35).

Second, centrality matters because it is linked to the ability to change and to how widely spread out the consequences of such change will be. When a personality component is central, it is likely to be dependent on many other components (and vice versa), so it will be more difficult to change. Changing a habit of not fulfilling promises, for example, is more likely to be difficult because in order to change that component, there are many others that may need to be changed as well (e.g., sympathize more with other people's needs and learning to finish things). Drawing analogy to a trade network, there are tradesmen who operate as pivotal points in the network (i.e., as central nodes): they have a large influence on the total productivity of the network (how much money the network as a whole makes), and it is very difficult to drive them out of business because of their strong connections with so many others (i.e., individual components of personality). It is unlikely but for some reason, it might be that a central component in fact does change (in the trade analogy, a pivotal person goes out of business). If so, then the consequences for the remainder of the network will be more widespread than if a peripheral component (tradesman on the periphery) changes. For instance, if one ceases to take voting seriously, this is not likely to have major effects on other aspects of one's expression of personality. In contrast, if because of whatever circumstance, one ceases to be a reliable person—as might occur in the early phases of dementia with a deterioration of memory functions—this is likely to have effects throughout the system.

Network structure as a source of stability

Human actions are flexible and unpredictable across situations, but at the same time, general patterns of behavior can be extremely rigid and very difficult to change. Theories of personality aim to reconcile these two facts of human life and provide compelling explanations for the stability that apparently underlies the great variability in daily moods, thoughts and behaviors. The traditional way of dealing with this issue is to invoke a two-part explanation in which the variation in behavior is governed by transient factors, whereas the average around which these variations are dispersed is caused by a stable factor. The latter is typically conceptualized as a trait, defined as a relatively enduring organismic (psychological, psychobiological) structure underlying an extended family of behavioral dispositions (Tellegen, 1991). Thus, in this definition, a 'trait' is a 'common cause', a structure that 'explains' the stable level of functioning around which a certain variability in 'states' revolves (e.g., the trait-state-error model; see Kenny &

[3] For the sake of simplicity, this chapter focuses on activation as a dichotomous characteristic of a personality component: it is either "on" or "off". However, it is also possible to define activation as an ordinal or continuous characteristic, in which components' activation varies along a scale. In this way, components in a personality network can be active with a certain 'intensity'.

Zautra, 1995). An example of such a structure is the latent dimension of extraversion, which is thought to cause stability by affecting the chances for a broad range of states to occur, as shown in the left panel of Figure 6.3.

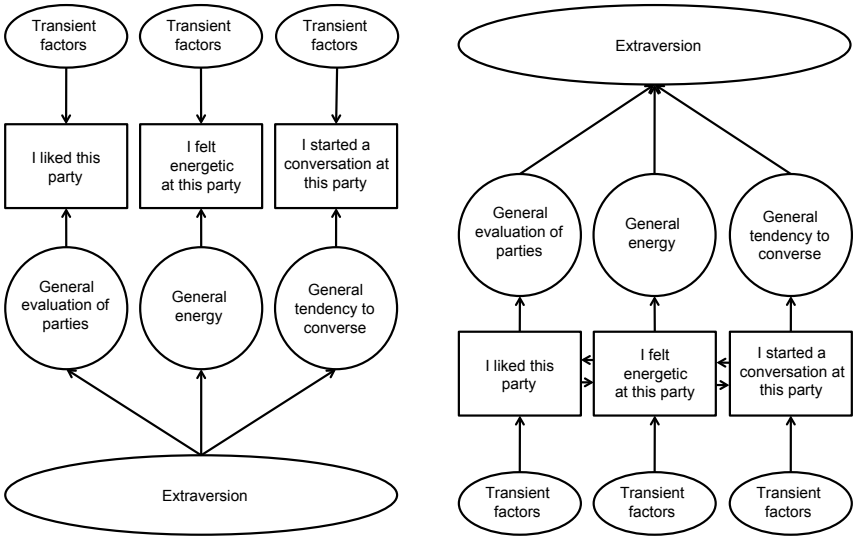

Figure 6.3: Illustration of the trait view according to a traditional latent variable (left panel) and a network perspective on personality (right panel). From a latent variable perspective, a trait such as extraversion is a common cause of stable dispositions that, together with transient factors, explain momentary states. The network alternative views direct interactions between personality components, influenced by transient factors, as the source of synchronized stability of components. In this view, a trait such as extraversion emerges out of these interactions. Traits are no longer common causes but summary statistics or index variables describing the average activation level of states.

Both the traditional model in the left panel of Figure 6.3 and the network in Figure 6.2 are between-subjects models that—for several reasons—cannot be assumed automatically to generalize to specific individuals (Borsboom et al., 2003; Borsboom, Kievit, Cervone, & Hood, 2009). From a network perspective, inference at the level of the individual is possible if one assumes that the dynamic structure of personality components of an individual can be represented in a similar network form (e.g., Figure 6.4 is an example of such a hypothetical network of an individual). Individual differences can then be captured by allowing for individual differences in components and the strengths of the connections among them.

From a network perspective, there are multiple ways in which trait-like and state-like characteristics can be defined at the level of individual networks (see Figure 6.4 for an illustration). This flexibility stands in stark contrast to the trait view, in which traits and states can only be sensibly defined at the level of the (first-order or second-order) latent dimensions (from a latent trait perspective, it would make no sense to define states and traits at the level of the items, although it might be technically possible). As such, individual differences can only be expressed in terms of that latent dimension as in, for example, 'Alice is more trait neurotic than Bob', whereas the network perspective can express many differences between Alice and Bob, such as 'Alice's network has more trait-like neurotic components than Bob's' and 'The connections in Alice's network are more state-like than Bob's'. Such multiple observations are likely more true to the subtle

nature of individual differences.

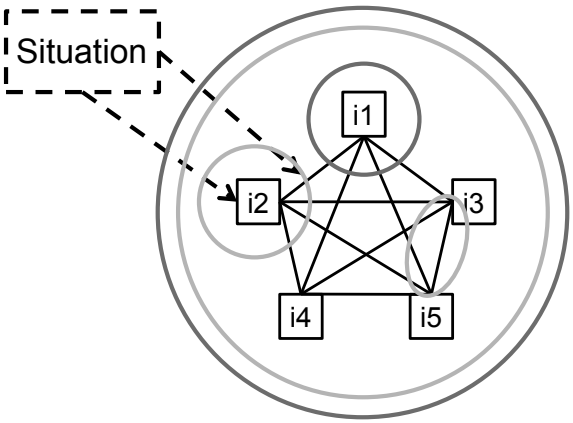

Figure 6.4: The possibilities of conceptualizing traits and states within an individual's network of five personality components (i1-i5). The light grey circles refer to possible conceptualizations of traits, whereas the medium grey circles refer to possible conceptualizations of states in the network. Situations in the environment can influence either individual components or connections among them, thereby changing their states.

A first way to define traits and states in a network is at the level of the network as a whole (turquoise circle around the entire network in Figure 6.4): synchronized stability of multiple components can result in the emergence of a stable trait such as extraversion (as illustrated in the right panel of Figure 6.3, in this case for extraversion components). That is, instead of the current view that a trait ultimately 'causes' behavior (i.e., arrows pointing from Extraversion in the left panel of Figure 6.3), the network perspective views a trait as a phenomenon that is the 'result' of (and, in that sense, emerging from) direct interactions between thoughts, feelings and behaviors as measured with personality items (i.e., arrows pointing towards Extraversion in the right panel of Figure 6.4). As such, a trait, from a network perspective, is similar to a summary statistic or index variable that describes the average activation level of states, which is consistent with the key assumption of a formative model (Edwards & Bagozzi, 2000). Importantly, the network perspective is thus *not* contradictory to current trait theories. Networks result in traits too, the only difference with current trait theories being that in the latter case, traits are most typically explained with a latent variable model in mind. That is, trait theories are currently intertwined with a latent variable perspective (as depicted in the left panel of Figure 6.3).

Networks can result in traits because transient factors and context determine the activation of affective, cognitive and behavioral components that in turn may activate one another if they are connected in the network architecture. Every time a set of components is activated (e.g., when a person feels energetic), the activation contributes to a self-evaluation stored in memory ("I am an energetic person") that is of the kind queried in typical personality questionnaires ("Would you consider yourself an energetic person?") and serves as evidence for evaluating the self-related hypothesis (van der Maas, Molenaar, Maris, Kievit, & Borsboom, 2011). General evaluations that arise from densely connected areas in the network will covary; as a result, these variables will form a large principal component if submitted to a data reduction technique such as principal components analysis. However, a simple structure confirmatory factor model (like that in the left panel of Figure 6.3) may not fit well because of violations of conditional independence

(i.e., because the model does not get the causal structure right). We understand this to be typical in personality research where confirmatory models can fit badly even though principal component structures are robust and replicable (McCrae, Zonderman, Costa, Bond, & Paunonen, 1996).

A second way in which networks can display trait-like and state-like properties in individuals' networks is at the level of these individual components themselves (turquoise and pink circles around the item boxes in Figure 6.4). Roughly speaking, there are two reasons why components can display both trait-like and/or state-like features. First, the wording of the items themselves may or may not refer to stable behavioral dispositions (e.g., "I easily feel offended by other people" and "I finish things once I have started them"), whereas the responses to other items may greatly vary over time (e.g, "I feel offended now" or the items represented in Figure 6.5). The latter components can be considered to be inherently more state-like, whereas the first are inherently more trait-like. Second, the activation of components can be altered (from 'active' to 'not active' or vice versa), depending on a specific 'situation' a person is in (orange arrow from situation to i2 in Figure 6.4). Some of these components are more state-like because alterations in the environment (i.e., different situations) result in unstable activity patterns (i.e., the change in activity is relatively temporary). For example, a component such as "I'm full of ideas" can be unstable in certain people: the component would be active (i.e., Alice feels full of ideas) for Alice after a positive day at work during which her boss complimented her on having a good idea but inactive (i.e., Alice does not feel full of ideas) the next day because her mother-in-law describes her in her face as a follower and not a leader. In contrast, some situations result in long-term stable changed activity in one or more nodes. For example, Bob, a trusting person, obtains a venereal disease from his cheating girlfriend who also dumps him. Subsequently, Bob re-examines basic assumptions about how he sees the world and as a result, changes: becomes less trusting, more suspicious of the motivations of others and so on.

Situations can also influence the connections among the components (orange arrow from situation to the connection between i1 and i2 in Figure 6.4; analogous to moderation). Connections subject to such influences can be more susceptible to change and thus more state-like in that they are aspects of personality that vary in response to different situations (analogous to what is hypothesized in the CAPS model: Mischel & Shoda, 1995, 1998). For example, Bob normally does not feel guilty because he sometimes feels just miserable for no reason (i.e., relatively stable weak connection between feeling miserable and feeling guilty). But, when Bob feels just miserable right when his wife surprises him with tickets for a cruise, he feels incredibly guilty: that is, the connection between feeling miserable and feeling guilty is stronger, triggered by the situation. It is, on a related note, this very malleability of certain connections that is the focus of many psychological treatment strategies (e.g., cognitive behavioral therapy; see Cramer et al., 2010). Other connections are likely relatively trait-like, in part, because the components they connect are inherently more trait-like as well, for example, the connection between "I like to go to parties" and "I feel comfortable around people".

The empirical study of this dynamic structure of personality networks becomes possible through the use of time series data. For instance, Figure 6.5 presents empirical correlation networks of four people who participated in a larger study into the effects of mindfulness training on a range of emotion and psychopathology variables (Geschwind, Peeters, Drukker, van Os, & Wichers, 2011; see supplementary materials available at http://www.aojcramer.com for a description of the sample and the measures). The participants in this study were assessed multiple times a day by using an experience sampling protocol, which generates series of observations over time. The networks in Figure 6.5 represent the lag-1 correlations between time series of four variables: anxiety, feeling

down, irritability and the pleasantness of the event reported to be the most important one during the assessment period. Specifically, a thick green arrow from A to I means that a higher level of anxiety at t predicts a higher level of irritability at $t + 1$; a thick red arrow from E to A means that a more positive evaluation of the event that took place at t predicts a lower score on anxiety at $t + 1$ and so on. Naturally, it is also possible to construct such intra-individual networks for correlations within the same time frame: the construction and interpretation of such graphs would be analogous to the procedure explicated for the inter-individual network that was presented in Figure 6.2.

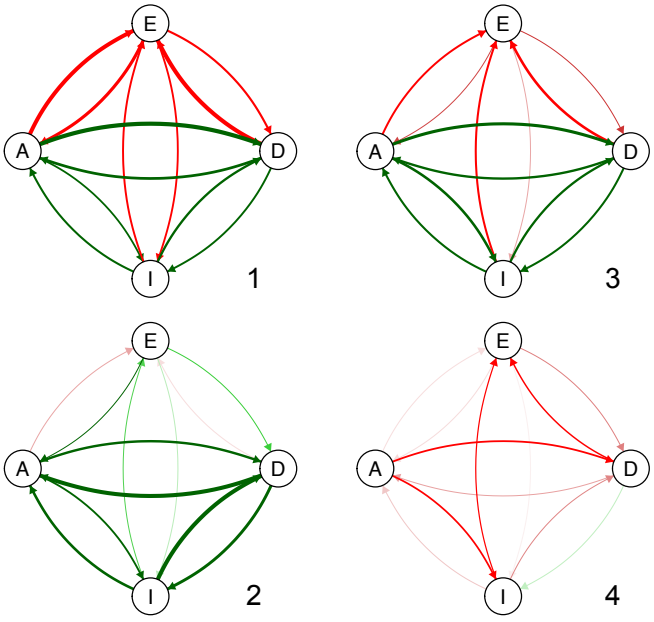

Figure 6.5: Network representations of the temporal dynamics of four individuals (1, 2, 3 and 4) who were repeatedly assessed in an experience sampling study. An arrow from, for example, node A to node I represents the correlation between the score on node A at time t with the score on node I at time $t + 1$: green (red) lines represent positive (negative) correlations. The thicker the arrow, the stronger the connection. E: pleasantness of the event reported to be most important; A: anxious; D: feeling down; I: irritable.

The individuals showed marked differences in their dynamic structure. Individual 1, who is relatively typical for the sample studied here, showed positive dependencies among A, I and D: for example, the more anxious at t, the more irritable at $t + 1$ (and vice versa). A, I and D all had negative dependencies with E: for example, lower anxiety at t predicted higher pleasantness of the event reported at $t + 1$, but a more pleasant event at t also predicted lower anxiety at $t + 1$. This appears not to be the case for individual 2 whose relations between E and the psychological variables were one-way traffic: for example, lower anxiety at t predicted higher pleasantness of the event at $t + 1$, but a more pleasant event at t did not appear to predict lower anxiety at $t + 1$. One might speculate that this individual 'profits' less from positive events. Individual 3 also showed this pattern but in addition showed no noticeable predictive relation between the psychological variables at t and the pleasantness of the event at $t + 1$. This individual thus appeared to function independently of the events reported in the relevant time. Finally, individual 4 showed a surprising pattern of purely negative relations, in which

the anxiety variable functioned as a source node without substantial incoming effects and seems to steer the other variables in a counterintuitive way ('increased' anxiety at t predicted 'decreased' irritability and depressed mood at $t + 1$, whereas 'decreased' irritability and depressed mood at t predicted a 'more' pleasant event at $t + 1$). We do not know, from the present data, to what extent these patterns generalize outside the studied time window or whether they have meaningful connections to the everyday functioning of the studied individuals. However, the differences between the network structures are quite suggestive and may, in future research, be shown to have significant consequences.

Thus, from a network perspective, the components of individuals' personality networks as well as the connections among them can exhibit trait and/or state-like properties, in part influenced by situations that figure as separate nodes in the network (see also Figure 6.4), and traits such as extraversion, or openness, emerge out of the combined activity of the components of the personality network, instead of being the common cause of these components. Traits as emerging entities do not violate some definitions of traits: for instance, the definition of Tellegen (1991) of traits as "enduring [...] structure[s] underlying an extended family of behavioral dispositions" would in fact seem neutral on whether the structure in question is a latent dimension or a network structure.

Understood in this way, the network perspective offers a possible resolution between trait approaches and situationist approaches that emphasize that traits can be adequately described as situation-relevant reaction patterns (e.g., Mischel & Shoda, 1995): the connections among situational nodes—external to the human system—and components that are more internal to the human system are likely to differ in strength across individuals. Such differences in situation-behavior associations lead to if-then signatures of the kind identified by Mischel and Shoda (1995).

Given the ample opportunities for individual differences to arise in a personality network structure, is it in fact possible that both stable individual differences and significant day-to-day variation arise from the same network structure? The answer is yes. To illustrate this, Figure 6.6 shows three simple networks, representing three fictitious people, consisting of three binary nodes (i.e., nodes that can either be 'active': 1 or 'inactive': 0). All variables are measured at multiple time points but without implying any direction of causation (i.e., all variables influence all other variables). Thus, each resulting network is an intra-individual representation of how three variables influence one another bidirectionally over time. The only differences among these networks are the strengths of the connections among the nodes (i.e., each connection has a certain weight that determines its strength): the rightmost network in Figure 6.6 is the most strongly connected, whereas the leftmost network is the least strongly connected. At time point t, whether or not a node is active is dependent on the status (0 or 1) of each of its neighbors times the relevant connection weight, which results in a total incoming effect A. The probability that the node is active then depends on the total incoming activation as follows:

$$P(X^i_{t+1}) = 1/(1 + e^{A}_{t}), \text{[4]} \tag{6.1}$$

n = number of nodes in the network;
X^{i} = node i;
i = 1, 2, ..., n;
t = time;
$A = n * n$ activation matrix

[4] Please note the similarity of this equation to equations in item response theory (e.g., Lord, 1953)

If we simulate data points according to this model for the three networks in Figure 6.6, the networks will all transition between activation patterns in a random fashion. That is, there will be significant ('day-to-day') variation in which nodes are active or inactive. On the other hand, the probability distributions of the total activation scores (i.e., the total number of nodes that are active at a randomly chosen time point for each network) will be stable: in Figure 6.6, the average activation level of the leftmost network will be lowest, whereas that of the rightmost network will be highest, and this is no surprise, given the fact that those networks are weakly and strongly connected, respectively. Thus, stable individual differences in average activation levels are possible as well, and it is exactly that synchronized activity of consistent patterns of node activation within individuals that may give rise to traits: if Figure 6.6 represented openness networks, then the person with the rightmost network would likely be an open person—because, on average, many openness nodes are active at the same time—whereas the person with the leftmost network would likely not be an open person. So given this potential of individual differences in network structure to generate both traits and day-to-day variations without invoking any latent dimensions, what, in turn, could cause these differences in network structure?

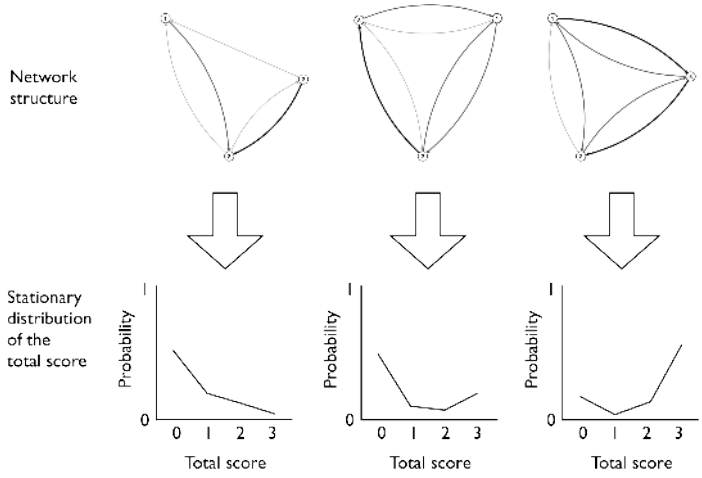

Figure 6.6: How network structures can lead to stability. The three network structures at the top of the figure, each representing a fictional individual, differ in connection strength: the darker a connection between two nodes, the stronger that connection. These structures generate stationary distributions at the bottom panel. The stationary distribution depicts the probability (y-axis) that, at a randomly chosen time point, a given number of components (x-axis) is active. These distributions are themselves stable over time, so if the number of active nodes were measured at repeated time points, the resulting scores would show high test-retest correlations.

The genetic origins of personality networks

Genes influence many human characteristics, and personality is one of them. Multiple studies have shown that personality dimensions are at least moderately heritable (Boomsma et al., 2002; Bouchard, 1994; Jang, Livesley, & Vernon, 1996; Kendler & Myers, 2010; Riemann, Angleitner, & Strelau, 1997; Loehlin, 1992): for example, 40% of the phenotypic variance in extraversion can be explained by additive genetic factors.

Assigning one number to represent heritability of any particular personality dimension makes sense from a latent trait perspective: items are no more than indicators of a common underlying trait (e.g., extraversion). As such, what is transmitted via genes from one generation to the next is the predisposition for developing that personality trait, not the propensity for a particular type of behavior/emotion/cognition as measured with a single item (i.e., personality component; see the left panel of Figure 6.7). In analogy with height, height is a latent trait (i.e., height is an unobserved variable for which we need measurement instruments to quantify it in individuals; see Bollen, 2002; Borsboom, 2008a)[5], which is measured with various methods (e.g., measurement tape). What is heritable is relative height itself—that is, children of tall parents tend to be tall as well—not any particular measurement of height.

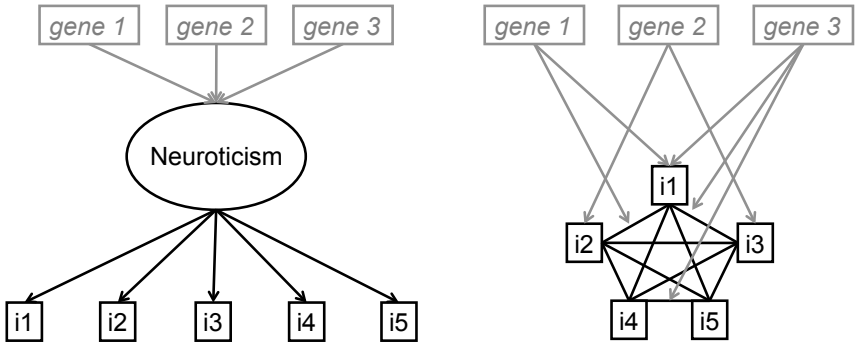

Figure 6.7: The influence of genes (grey boxes) on neuroticism according to a latent trait perspective (left panel) and a network perspective (right panel) on personality. Left panel: genes influence the individual items (i1-i5) not directly, only indirectly via the latent trait 'Neuroticism'. Right panel: genes influence the individual items and connections between them directly.

One can go one step beyond defining heritability as a characteristic of a personality dimension such as extraversion (whether it be a latent trait or an emerging feature, as it is defined in a network model), namely, by defining heritability as genetic influence on the individual components of the network and the connections between these components (see the right panel of Figure 6.7). For example, it might be that liking parties is 20% heritable, whereas enjoying the company of other people is 65% heritable, the difference not being due to differences in reliability (consistent with a latent independent pathways model). Likewise, it could make sense to say that the degree to which people who are quick to understand things have a tendency to be full of ideas (i.e., connection strength between these components) is 34% heritable or the degree to which people who regularly just feel miserable have a tendency for suicidal ideation (and vice versa) is 78% heritable.

Now, if the network perspective is accurate in portraying personality, then current techniques for the next step in behavioral genetic research, that is, the identification of genes that are the driving forces behind these heritability estimates, might prove problematic. Current techniques employed in genetic association studies typically rely

[5] Some may argue that height is not an appropriate example of a latent variable. However, see these papers for an explication of the point that latent variables are variables that are not directly observed (i.e., we cannot observe directly whether someone is 5.1 ft or 5.2 ft). We therefore need measurement instruments to quantify such variables. Also, for example, Harman (1960) argued that a latent variable is an underlying variable that helps explain why certain other variables correlate (i.e., two methods to measure height in Bob correlate because they are caused by the same underlying variable, namely Bob's height). In both non-formal views of the theoretical status of a latent variable, height is an appropriate example.

on a sum score (e.g., the sum of the neuroticism item scores of the NEO-PI) as a proxy for the latent variable, neuroticism in this example. Genetic association studies in their most rudimentary form identify genes or genetic variants as being associated with a particular personality dimension if they predict the sum score (i.e., the dependent variable in the design: Cramer, Kendler, & Borsboom, 2011; van der Sluis, Kan, & Dolan, 2010). If personality dimensions were indeed latent traits, this approach is sensible, although not necessarily optimal (van der Sluis, Verhage, Posthuma, & Dolan, 2010).

To date, standard genetic linkage and association studies have not yielded any clear genetic candidates: for the Big Five personality dimensions, many candidate gene findings are not replicated, and the genetic polymorphisms that are consistently identified typically account for less than 2% of the genetic variance (Amin et al., 2011; de Moor et al., 2010; Fullerton et al., 2003; Kuo et al., 2007; Nash et al., 2004; Terracciano et al., 2010). This discrepancy between moderately high estimates of population heritability and the inability to identify the responsible genetic polymorphisms is called the 'missing heritability' problem, a problem that is pervasive throughout the entire realm of psychology as well as other complex biomedical traits such as height and blood pressure (Maher, 2008; Manolio et al., 2009).

Although many explanations have been put forward for the missing heritability problem (e.g., additive small effects of many individual genes, limited sample size, population stratification and selection bias; K. A. Frazer, Murray, Schork, & Topo, 2009; Maher, 2008; P. Sullivan, 2011), we focus on another possible reason: misconceptualization of the phenotypic model (6.7). In particular, the model in the left panel of Figure 6.7 might be wrong (as was, for example, recently shown for nicotine dependence where two genes influenced individual symptoms quite differently: Maes et al., 2011). From a network perspective (the right panel of Figure 6.7), it is not likely that all components and connections between them in the personality network are influenced by the exact same set of genes: gene 1 influences different parts of the network than gene 2. For example, components such as feeling sad and finding political discussions boring probably involve different antecedent pathways: feeling sad has more to do with emotional processes whereas finding political discussions boring is more likely a cognitive phenomenon, and as such, feeling sad and finding political discussions boring probably involve different biological substrates and pathways and thus different genes. If so, then attempting to relate genetic polymorphisms to their sum score is not likely to contribute to effective gene hunting because with a sum score, one only captures the genetic variance that is shared among the components and their connections (van der Sluis, Verhage, et al., 2010): the power to detect effects from single-nucleotide polymorphisms (SNPs) in sum scores is multiple times lower when these gene effects are local (e.g., gene 2 in the right panel of Figure 6.7 influences two neuroticism items) compared with when these effects are global (e.g., gene 1 in the left panel of Figure 6.7 influences all neuroticism items via the latent trait 'neuroticism').

It is hard to pit the models in Figure 6.7 directly against each other because estimation and fitting algorithms for the network model have not been developed in sufficient detail. However, we can examine and test divergent predictions of the models such as the location of the effect of SNPs: from a latent trait perspective, one would expect SNPs to impact at the latent trait level, whereas from a network perspective, one expects SNPs to impact at the level of the individual components (Maes et al., 2011).

We tested this prediction by using data from 1625 healthy individuals who participated in the dbGAP GAIN Major Depression Disorder study (dbGAP study accession, phs000020.v2.p1). In particular, we investigated the effects of seven top SNPs that were implicated in neuroticism in two recent genome-wide association studies (de Moor et al., 2010; Terracciano et al., 2010). We tested whether the effects of these genes on the item

responses were most likely to be mediated by the latent trait 'neuroticism' or whether these effects were more likely to be item specific (see supplementary materials available at htpp://www.aojcramer.com for an extended description of the sample and the method).

The analyses showed that in this sample, none of the seven top SNPs had a significant direct influence on the latent trait 'neuroticism'. On the one hand, this result can be interpreted as a non-replication of these SNPs in this sample, which could be due to a limited sample size or the use of a different neuroticism instrument (see supplementary materials available at http://www.aojcramer.com). On the other hand, the result can be interpreted as lack of support for a latent trait perspective on the influence of genes on personality dimensions. At the same time and consistent with a network perspective, we did find evidence for significant direct influences of three SNPs (rs17453815, rs12509930 and rs7329003) on three individual neuroticism items ("restless", "can't sit still", "guilty" and "sleepless due to thought racing": see Figure 6.8). These effects were significant at $\alpha = 0.005$ (p-values for the SNP-latent trait 'neuroticism' relations ranged between 0.16 and 0.62; see supplementary materials available at http://www.aojcramer.com). Naturally, replication of these specific SNP-item relations in other, larger samples is imperative to draw definitive conclusions. This example with real data mainly serves to illustrate how to test the diverging predictions from the latent variable versus the network perspective.

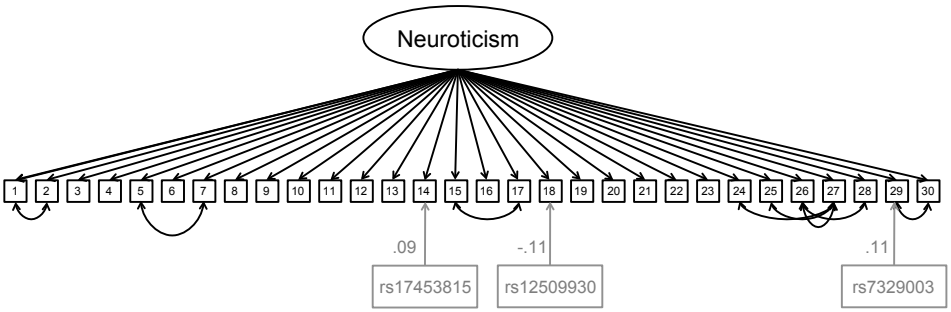

Figure 6.8: The phenotypic latent variable model relating the latent neuroticism factor (black oval) to 30 neuroticism items (black rectangles). In light grey are the significant relations of three single-nucleotide polymorphisms (SNPs) (light grey boxes) with individual items. Relations between the SNPs and the latent neuroticism factor were not significant. Two-sided arrows represent correlations; one-sided arrows represent regressions. The impacts of the SNPs on the neuroticism items are expressed as standardized regression weights. *1*: fantasy; *2*: daydreams; *3*: lonely; *4*: undecidedness; *5*: wandering thoughts; *6*: nervous superiors; *7*: distracted; *8*: not old self; *9*: nightmares; *10*: listless; *11*: up/down energy; *12*: senseless thoughts; *13*: worry about set back; *14*: restless, can't sit still; *15*: miserable; *16*: needs to be alone; *17*: up/down spirit; *18*: guilty; *19*: handle disappointment; *20*: take things at heart; *21*: easily hurt; *22*: nervous; *23*: worry about past; *24*: dissatisfied; *25*: irritable; *26*: even tempered; *27*: sad; *28*: mood swings; *29*: sleepless due to racing thoughts; *30*: sleepless worry.

Another way of testing the viability of the latent trait perspective is to check whether the directions of the effects of the seven top SNPs (i.e., increase or decrease risk) are the same across the individual neuroticism items. If the latent trait perspective is correct and genes influence individual items only indirectly via the latent trait 'neuroticism', then all relations between a SNP and identically coded neuroticism items should have the same sign. This, however, is not what we found in this data set (see supplementary materials available at http://www.aojcramer.com for the odds ratios (OR) between SNPs and neuroticism items and p-values computed according to false discovery rate criteria). For example, SNP rs17453815 was associated with a decreased risk for being "easily irritable" (OR = 0.69, $0.01 < p < 0.05$) but with an increased risk for "restless, can't sit

still" (OR = 1.32, $0.001 < p < 0.01$). Similarly, SNP rs11707952 was associated with a decreased risk for "experiencing mood swings" (OR = 0.73, $0.01 < p < 0.05$) with also with an increased risk for "not feeling your old self" (OR = 1.22, $0.01 < p < 0.05$).

Given that the current methodological state of affairs does not allow for a direct statistical test, psychometric modeling with genetic data might provide a fruitful avenue to explore the feasibility of latent trait versus network models because these models come with specific predictions that can be tested in a confirmatory factor analytic framework. As such, we do not take the aforementioned results to signify anything definitive about SNP effects on neuroticism items; rather, these results serve as concrete examples of how one might go about testing predictions of the two competing models. It might be argued that finding local effects (i.e., SNP effects on individual items) is not in violation of the 'statistical' aspects of the latent trait model. Although this is true, the 'theoretical' notion of a latent variable as an accurate reflection of personality dimensions is much harder to maintain in the face of genetic effects whose impact is not at that latent level but, instead, at the item level.

If future evidence favors the network model, the next step would be to wonder how personality networks are tied to psychopathological phenomena. As we have argued elsewhere (Borsboom, 2008b; Cramer, 2012; Cramer et al., 2011, 2010), mental disorders can also be understood in terms of networks of interacting symptoms (e.g., insomnia $\rightarrow$ fatigue $\rightarrow$ concentration problems). Because it is well known that certain personality dimensions predict the development of certain forms of psychopathology (Hettema, Neale, Myers, Prescott, & Kendler, 2006; Kendler, 2006; Terracciano, Lockenhoff, Crum, Bienvenu, & Costa, 2008; van Os & Jones, 2001), how might this covariation arise from a network perspective?

The roads from personality dimensions to mental disorders

Some aspects of personality are correlated with the onset and/or maintenance of certain mental disorders: for example, (1) trait neuroticism and major depression (MD), (2) alienation (a tendency to feel mistreated, victimized, betrayed and the target of false rumors) and substance dependence and (3) high negative emotionality (a propensity to experience aversive affective states) and antisocial personality disorder (D. N. Klein, Kotov, & Bufferd, 2011; Krueger, 1999; Krueger, Caspi, Moffitt, Silva, & McGee, 1996). From a latent variable perspective—in which a personality dimension and a mental disorder are latent entities—there are three ways in which personality features (P) and mental disorders (M) can be modeled (see Figure 6.9): (1) models in which P and M are not causally related in whatever shape or form. Instead, P and M are correlated because they are (partly) influenced by the same etiological processes (the A arrows in Figure 6.9); (2) models in which P is an effect of M (the B arrow in Figure 6.9); and (3) models in which P precedes M (the C arrow in Figure 6.9).

From a network perspective, the three classes of models as depicted in Figure 6.9 do not work because in both personality and mental disorder networks, there are no latent variables. Because items are at the heart of personality networks and symptoms at the heart of mental disorder networks, the most sensible way to conceive of relations between the two networks is by means of direct relations between these items and symptoms (see blue lines in Figure 6.10). Instead of one option for three types of pathways between personality and psychopathology (i.e., A, B and C model in Figure 6.9), each blue line between an item and a symptom in Figure 6.10 represents an optional pathway that can be of the A, B or C type. For example, in Figure 6.10 (without implying causality because

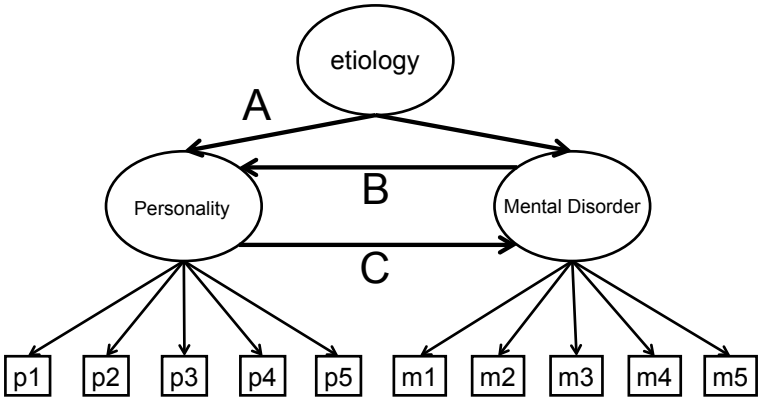

Figure 6.9: Three ways of modeling the relationship between a personality dimension and a mental disorder. In all three models, personality dimensions (P) and mental disorders (M) are hypothesized to be latent variables (ovals) that have causal influence on the items that are used to measure these variables (p1-p5 for personality and m1-m5 for mental disorder). The models hypothesize that either (A) P and M are related via common etiological processes, (B) P is an effect of M or (C) precedes M.

there are no arrows in the figure), one pathway from personality to mental disorder (and vice versa) could be: p2 - p4 - m5 or, alternatively, m2 - p3 - p5. That is, from a network perspective, pathways between items and symptoms indicate dependencies between them, such that one may activate another, analogous to how diseases spread through a population. For example, the tendency to feel nervous around other people (p2) likely increases the probability of spending much time alone (p4), which may result in relatively frequent feelings of anhedonia (m5). The other way around may be an equally likely pathway: prolonged feelings of anhedonia may well undermine the capacity to enjoy the company of other people.

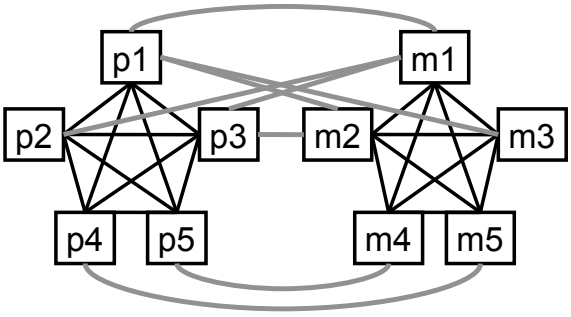

Figure 6.10: Modeling the relations between personality dimensions and mental disorders. Items from a certain personality dimension (p1-p5) that are connected with one another (black lines) are directly connected (grey lines) with symptoms of a certain mental disorder (m1-m5) that are also connected with one another (black lines).

As a starting point, like in Figure 6.2, correlations between personality items and mental disorder symptoms could be used as quantifications of the strength of the connections between these items and symptoms. Figure 6.11 shows such a correlation network for neuroticism and MD data obtained from the Virginia Adult Twin Study of Psychiatric

and Substance Use Disorders (Kendler, 2006; Prescott et al., 2000) (see supplementary materials available at http://www.aojcramer.com for a description of the sample and the measures). Some marked differences in connection strength among the items and symptoms stand out. First, there are clearly two clusters of strongly connected items/symptoms, one corresponding to neuroticism (blue nodes) and the other corresponding to MD (red nodes). Second, some neuroticism items are more strongly connected to MD symptoms than other neuroticism items (and vice versa): for example, feelings of worthlessness (*wort*: MD) and feelings of loneliness (*lone*: neuroticism) are more strongly connected than one's feelings being easily hurt (*hurt*: neuroticism) and increased appetite (*iapp*: MD).

Within a network such as the one in Figure 6.11, central nodes might be the crucial nodes on pathways connecting neuroticism and MD because such nodes are strongly connected with both neuroticism and MD nodes (as argued earlier in the chapter). So for example, in this particular sample, feeling just miserable (*mise*: neuroticism), being a nervous person (*nerv*: neuroticism), feelings of loneliness (*lone*: neuroticism) and feelings of worthlessness (*wort*: MD) are the most likely candidates for being part of the multiple pathways from neuroticism to MD (and the other way around).

Another way of generating hypotheses about likely pathways from personality to psychopathology (and vice versa) is through partial correlations. The general idea is the same as with simple correlations—one constructs a network with the strengths of the connections between the nodes reflecting the magnitude of the correlations—but partial correlations are potentially more informative about whether two variables are in fact truly related. A high simple correlation between two variables does not necessarily imply that a unique relation exists between these variables. For instance, a high correlation between feelings of guilt and feelings of worthlessness may be due to the fact that both components are influenced by another component in the network, for example, depressed mood. As such, feelings of guilt and feelings of worthlessness are not uniquely related; the correlation arises because of their common cause, depressed mood. If that is true, the correlation between feelings of guilt and feelings of worthlessness should be (very) low when depressed mood is controlled for, and this is exactly what a partial correlation does: it quantifies the association between any two components while controlling for one or multiple other components in the network. As such, when one computes correlations among the neuroticism and MD items/symptoms while controlling for all other components in the network, a high partial correlation is potentially more indicative of a true relation than a simple correlation. Figure 6.12 presents such a partial correlation network on the basis of the same data that was used for Figure 6.11.

A few things stand out when inspecting Figure 6.12. First, many connections are weaker in Figure 6.12 compared with connections between the same components in Figure 6.11, for example, the connection between feelings of worthlessness (*wort*: MD) and feelings of loneliness (*lone*: neuroticism): a direct relation between these components might exist (the partial correlation is not close to 0) but is likely partially influenced by other components in the network (because the partial correlation is lower than the simple correlation). On the other hand, feelings of worthlessness (*wort*: MD) and feelings of guilt (*guil*: neuroticism) are almost as strongly connected in both figures: these two components are likely directly related without being substantially influenced by other components in the network. Second, some pathways from neuroticism to MD (and vice versa) are more likely than others (i.e., are more strongly connected compared with other pathways): for example, a pathway via feelings of worthlessness (*wort*: MD) and guilt (*guil*: neuroticism) is more likely than a pathway via weight loss (*wlos*: MD) and describing oneself as a nervous person (*nerp*: neuroticism). When considering which nodes are the most central in this network, the most likely candidates for playing pivotal

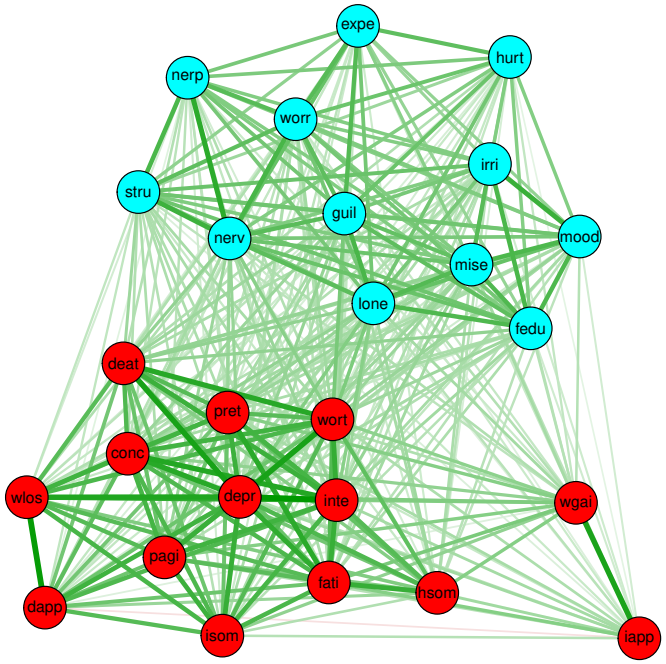

Figure 6.11: A network based on tetrachoric correlations between the 12 neuroticism items from the EPQ and the 14 disaggregated DSM-III-R symptoms of major depression (MD). The red nodes represent the individual MD symptoms, whereas the blue nodes represent the neuroticism items. Nodes are connected by green (red) lines if they are positively (negatively) correlated. The thicker the line, the higher the correlation. The same algorithm as in Figure 6.2 was used to generate the network: the most strongly connected nodes appear in the middle of the figure. *hsom*: hypersomnia; *isom*: insomnia; *fati*: fatigue; *iapp*: increased appetite; *dapp*: decreased appetite; *inte*: loss of interest; *conc*: concentration problems; *pret*: psychomotor retardation; *wgai*: weight gain; *wlos*: weight loss; *pagi*: psychomotor agitation; *depr*: depressed mood; *wort*: feelings of worthlessness; *deat*: thoughts of death; *mood*: mood often goes up and down; *mise*: sometimes "just miserable" for no reason; *irri*: irritable; *hurt*: feelings are easily hurt; *fedu*: often feels fed-up; *nerp*: rather nervous; *worr*: worrier; *stru*: tense or highly strung; *expe*: worries too long after an embarrassing experience; *nerv*: suffers from "nerves"; *lone*: often feels lonely; *guil*: often troubled about feelings of guilt.

roles in pathways from neuroticism to MD (and vice versa) are feelings of loneliness (*lone*: neuroticism), guilt (*guil*: neuroticism) and worthlessness (*wort*: MD); thoughts of death (*deat*: MD), being nervous (*nerv*: neuroticism) and describing oneself as a nervous person (*nerp*: neuroticism).

Partial correlations may be used to generate more parsimonious hypotheses about likely pathways from certain personality dimensions to certain mental disorders (and vice versa) but the technique is by no means bulletproof. It could be, for example, that a connection between two components with a low partial correlation does in fact exist. Sampling error, for example, might result in a low partial correlation between two components whereas a direct relation in fact exists in the whole population. Therefore, replication of findings in multiple samples is a necessity before any definitive conclusion can be drawn. Another way of testing hypotheses generated by partial correlations in between-subjects data is via longitudinal studies in which these hypotheses are verified in individuals. But instead of focusing on total scores on personality and psychopathology

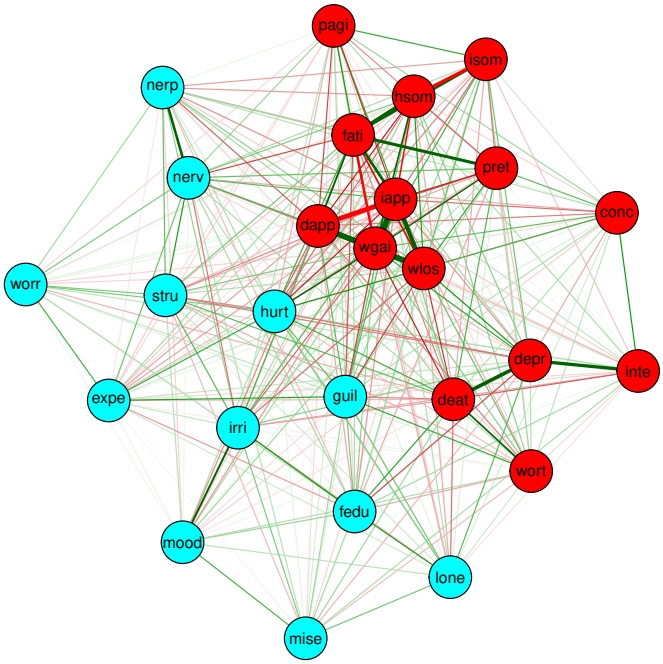

Figure 6.12: A network based on partial correlations between the 12 neuroticism items from the EPQ and the 14 disaggregated DSM-III-R symptoms of major depression (MD). The red nodes represent the individual MD symptoms, whereas the blue nodes represent the neuroticism items. Nodes are connected by green (red) lines if they are positively (negatively) correlated. The thicker the line, the higher the partial correlation. The same algorithm as in Figure 6.2 was used to generate the network: the most strongly connected nodes appear in the middle of the figure. *hsom*: hypersomnia; *isom*: insomnia; *fati*: fatigue; *iapp*: increased appetite; *dapp*: decreased appetite; *inte*: loss of interest; *conc*: concentration problems; *pret*: psychomotor retardation; *wgai*: weight gain; *wlos*: weight loss; *pagi*: psychomotor agitation; *depr*: depressed mood; *wort*: feelings of worthlessness; *deat*: thoughts of death; *mood*: mood often goes up and down; *mise*: sometimes "just miserable" for no reason; *irri*: irritable; *hurt*: feelings are easily hurt; *fedu*: often feels fed-up; *nerp*: rather nervous; *worr*: worrier; *stru*: tense or highly strung; *expe*: worries too long after an embarrassing experience; *nerv*: suffers from "nerves"; *lone*: often feels lonely; *guil*: often troubled about feelings of guilt.

questionnaires, which is the sensible thing to do when a unidimensional latent variable model holds[6] (see Grayson, 1988) —longitudinal studies from a network perspective would analyze each item and symptom separately for a prolonged time. In addition, with the time series techniques explicated earlier in this paper, the temporal pathways among these items and symptoms for individual people may be identified and directly modeled. Such studies undoubtedly will reveal many idiosyncracies—that is, there are likely many ways by which people develop certain forms of psychopathology as a result from certain personality characteristics (and vice versa)—but with some strong between-persons partial correlations, we have found in the data example earlier that some important commonalities can be expected as well. As such, the network perspective and its associ-

[6] In every unidimensional latent variable model, the sum score has a monotonic likelihood ratio with the latent variable, thereby rendering the sum score a better approximation of the latent variable than a single item.

ated investigation techniques may shed light on the exact nature of the complex relation between personality dimensions and mental disorders.

Conclusion

In this chapter, we have argued for a novel perspective on personality, in which the cognitive, affective and behavioral components of personality (e.g., liking parties and finding political discussions boring) are related through causal, homeostatic and logical connections. Traits such as extraversion and agreeableness emerge out of these connectivity structures, which implies a radical departure from traditional perspective in which traits are causes of the relevant components. We have shown how the network perspective may potentially alter our conception of what personality is and may supply new research techniques to investigate (1) overall personality architecture, (2) state and trait conceptualizations of personality, (3) the genetic background of personality architecture, and (4) the relations between personality and psychopathology.

Network methodology is far from fully developed. Examples concern the development of estimation and fitting algorithms for network models, robustness analyses for inferences on network structures, combining inter-individual and intra-individual data and model testing. Pertaining to the latter example, falsifying or confirming a network model can sometimes be quite complicated—for example, a unidimensional latent variable model will fit data that is generated by a network model in which all nodes are bidirectionally connected with equal strength—and sometimes surprisingly easy: for example, if one has the hypothesis that an inter-individual network is mutualistic (i.e., has only positive bidirectional connections so that nodes reinforce one another), then observing a negative correlation is enough to falsify that hypothesis. In its current state, network modeling could be compared to latent variable modeling in the 1950s: we have the ideas and the models but we still need to overcome many methodological obstacles. Nevertheless, the network perspective offers a plausible candidate model for explaining the 'common' structures of personality and the many idiosyncratic ways in which people deviate from that structure. One of its more attractive features is that the network perspective provides an intermediate position between traditional trait and situationist approaches, which both have longstanding traditions in personality psychology and which both have contributed greatly to our current understanding of personality. The network perspective takes the best of both worlds: it can explain how traits emerge out of the network structure, but it can also accommodate situational influences as external nodes that can activate individual components of the network (or connections among them).

Does adhering to the network perspective mean the end of factor analysis and other techniques associated with the more traditional perspectives on personality? No. Within the network perspective, factor analysis may become a useful technique for identifying groups of closely connected components. In fact, in special cases, it may be possible to estimate certain network parameters through factor analysis because groups of reciprocally connected components can behave exactly as predicted under a factor model (van der Maas et al., 2006). As such, we do *not* object to latent variable *modeling* in which conditional independencies implied by a statistical model are investigated and tested. Also, we readily acknowledge that some of the hypotheses that follow from the network perspective could in principle be tested with latent variable techniques (e.g., testing the influence of genes on individual personality items with independent pathway models) nor do we deny that if some relatively unexplored areas of the latent variable realm would be more extensively cultivated in personality research (e.g., intra-individual factor modeling over time and state-trait modeling within a latent variable framework; Steyer, Schmitt, & Eid, 1999), the latent variable model might be equally capable of accommodating

certain phenomena compared with the network perspective (e.g., accommodating both inter-individual differences and day-to-day variation).

The question of which techniques are capable of doing what is in our opinion not the one that should matter most in personality research. There is and should be no arms race at the level of the (future) technical accomplishments of both models. What matters most is which perspective provides the most plausible account of how personality arises: do traits cause cognitive, affective and behavioral components or do traits emerge from complex interactions between these components? How can future research help in finding an answer to this pivotal question? Given the current lack of methodological sophistication of the network models, the most likely frontrunner in terms of empirical research will be time series analysis of intra-individual data. Such data can, for example, be collected by assessing individuals' current thoughts, feelings and behaviors at many consecutive time points (for example by means of an experienced sampling protocol, which has been developed in considerable detail in clinical psychology). If time series analysis of such data would show that, within individuals, personality components have a (bi)directional influence on one another, then this would be strong evidence in favor of the network hypothesis and against the latent variable hypothesis. Another research strategy might be an inter-individual approach, in which one would experimentally test whether manipulating one personality component has an effect on another personality component.

In our view, the reification of factors such as extraversion as causes of individual behavior is unnecessary and unwarranted in the case of personality. That is, we do not object to latent variable *techniques* but we do object to a latent variable *theory* in which the measurement model with a common cause structure is interpreted as evidence for latent causal entities that operate in the minds of individuals, causing all sorts of cognitive, affective and behavioral patterns (see also Borsboom, 2008a). Human behavior simply does not appear to work this way: it is not extraversion that causes party going, liking people and enjoying conversations; it is liking parties, liking people and enjoying conversation, and interactions between them, that makes one an extraverted person.

Chapter 7

Measurable like temperature or mereological like flocking? On the nature of personality traits

Abstract

This chapter was written in response to a set of commentaries (abstracts of these commentaries available in Appendix B and full texts available at http://www.aojcramer.com), which were written in response to the previous chapter. We thank these commentators for their suggestions and critiques that have aided sculpting the ideas that are presented in this chapter. Some commentators wholeheartedly disagreed with the central tenet of the network perspective on personality, namely that traits are the result of mutual interactions between thoughts, feelings and behaviors. In this rejoinder, we primarily focus on these commentaries by 1) clarifying the main differences between the latent versus the network view on traits; 2) discussing some of the arguments in favor of the latent trait view that were put forward by these commentators; 3) comparing the capacity of both views to explain thoughts, feelings and behaviors. Some commentators were by and large positive about the network perspective, and we discuss their excellent suggestions for defining components and linking these to genes and other biological mechanisms. We conclude that no doors should be closed in the study of personality and that, as such, alternative theories should be welcomed, formalized and tested.

Adapted from: Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., Kendler, K. S., & Borsboom, D. (2012). Measurable like temperature or mereological like flocking? On the nature of personality traits. *European Journal of Personality, 26*, 451-459.

"No doors should be closed in the study of personality"
(Allport, 1946)

What are traits? To this question, there are almost as many answers as there are personality psychologists. The previous chapter was a first attempt at formulating a novel *theory* of personality (and not, as Schimmack and Gere suggest, merely a new analysis tool) in which traits do have a place. The difference with existing perspectives is that we do not see traits as *causes* of thoughts, feelings and behaviors (i.e., personality components)—the idea that has come to dominate personality psychology in the past decades—but as *consequence* of the interactions between such thoughts, feelings and behaviors. Thus, rather than reflective latent variables, personality traits are better conceived of as formative variables: summaries of relevant cognitive, affective and behavioral components that interact with one another in myriads of ways. We hypothesized that clusters of more strongly correlated components, typically interpreted as *signs* of underlying factors, in fact signal components that are particularly strongly interconnected. The coordinated behavior of these components thus *emerges* from the local interactions between them, just like flocking emerges from the local interactions between birds.

The commentaries we received are dividable in two general response categories: the first contains commentaries that were by and large positive, including very helpful suggestions for improving the precision and scope of the network perspective (**Costantini & Perugini; Denissen, Wood & Penke; Furr, Fleeson, Anderson & Arnold; Read & Miller**). The second class consists of commentaries that were (sometimes wholeheartedly) dismissive of our proposal, mainly because of reluctance to let go of the idea that personality traits are *necessarily* latent entities (e.g., **Guillaume-Hanes, Morse & Funder; Schimmack & Gere; Terracciano & McCrae**). The primary focus of this rejoinder pertains to this latter collection of commentaries. More specifically, we aim at (1) clarifying the main differences between the latent versus the network view on traits; (2) discussing some of the arguments in favor of the latent trait view that were put forward by commentators; and (3) comparing the capacity of both the network view and the latent trait view to explain thoughts, feelings and behaviors. Finally, we discuss some of the commentators' excellent suggestions for defining components and linking these to genes and other biological mechanisms.

Temperature versus flocking

Why do some aspects of personality, such as party-going behavior and liking people, cluster together? In the latent trait view, they do so because they are *caused* by the same underlying trait (extraversion in this example). This definition of a personality trait, as a cause of behaviors, thoughts and feelings, has come to permeate the field of personality psychology under many different names (source traits: Cattell, 1950; genotypic traits: Eysenck, 1967; trait$_2$: Wiggins, 1984) and is mathematically formalized in the generic latent variable model. In such a latent variable model, a personality trait and its items are associated with one another analogously to the relation between temperature and thermometers (Borsboom, 2008a): differences in temperature cause differences on thermometer readings via a well understood process by which particles exchange kinetic energy. If multiple thermometers are used (in the personality case: extraversion causes party-going behavior and liking people), thermometer readings are *measurements* of a common latent variable, namely temperature (extraversion is measured by items such as party-going behavior and liking people).

The hypothesized measurement relation between a trait and its items features promi-

nently in the five-factor theory of personality, most vehemently advocated by **Terracciano and McCrae**. Importantly, this and comparable factor models come with the assumption of *local independence* (Holland & Rosenbaum, 1986; Lord, 1953; McDonald, 1981): in the temperature analogy, a high correlation between the readings of two thermometers at the same time can be (almost perfectly, depending on the reliability of both thermometers) explained by the underlying cause of these readings, namely temperature. That is, if differences in temperature function as the *common cause* of differences in the thermometer readings, then there can be no direct relation between the two thermometers (i.e., a change in the reading on one thermometer does not change the reading on the other thermometer, and vice versa). In a measurement model, this is a highly sensible requirement. However, in our view, it is not a very plausible model for the relation between, say, conscientiousness and being in time for appointments.

This neither implies that factor models are useless nor that the results of factor analyses and related techniques cannot be reinterpreted along different lines. As such, **Ashton and Lee** are free to advocate their own non-causal definition of a factor as "a common element shared by its defining variables" (e.g., all birds have feathers), and we agree with them that non-causal interpretations of factors are compatible with a network perspective. For pragmatic reasons, it may also be useful to aggregate covarying individual differences into larger components and neglect the more stringent assumptions of factor analysis. However, we do not see how the assumptions of factor analysis sit with the idea that factors can be identified by *common elements*. The factor model does not hypothesize that there is a common element among indicators but that they share variance; moreover, that they share variance for a very special reason, that is, because they depend on the same latent variable. And *this* is what the psychometric model *is* consistent with, as the latent variable functions precisely as an unobserved common cause (e.g., having feathers is what causes certain animals to be birds; Pearl, 2000).

When adhering to a latent variable model-based explanation of the clustering of certain items, one cannot evade the local independence assumption: although it is technically not a problem to fit a one-factor model in which certain items are allowed to correlate, in addition to and independent of the relation that they share via the latent factor (i.e., to have a *direct* relation, a weaker form of the above-mentioned strict local independence assumption), the more such correlations are allowed to exist in the model, the less convincing is the case for an underlying trait that explains the majority of covariance between the items. Thus, when **Terracciano and McCrae** argue, in *defense* of the latent trait view, that liking parties is caused by both liking people and extraversion, they actually shoot themselves in the foot by admitting the existence of *direct* relations between the items of extraversion. If direct relations are allowed, factor analysis ceases to be a credible tool for identifying unobserved causes because that interpretation is crucially dependent on the assumption of local independence.

Naturally, other ways of tweaking this basic model of temperature are possible, and we acknowledge (again) that we fitted the simple model without distinguishing between first-order and second-order factors (e.g., **Ashton & Lee; Terracciano & McCrae**). However, **Terracciano and McCrae** are not committed to these more complex models either when they maintain their position that extraversion causes party-going: there is no distinction between first-order and second-order factors in this statement. Certainly, one may fit a much more complex model to the data with cross-loadings and lower order latent causes. However, in our view, this implies that to meet the assumption of local independence, one then introduces many extra, untested hypotheses. The alternative is to drop the unlikely assumption of local independence. As shown by van der Maas and colleagues (2006), a network of interdependent components can provide a valid alternative

for a well-fitting complex factor model. We stress that, in practice, factor analysis merely identifies clusters of items that correlate stronger with one another than with items outside the cluster. Hence, items that *load on the same factor* may be taken to identify networks of mutually reinforcing components (see van der Maas et al., 2006). In that case, a factor is not a latent trait with causal power but a summary statistic for how a set of items are influenced by one another (Cramer, 2012). This idea resembles Tellegen's (1991, p. 15) assertion that "A trait dimension is [...] a population concept representing an orderly statistical structure of covariation", as well as Mischel's (1973) thesis that traits are "summary terms [...] applied to observed behavior".

Pertaining to the network perspective, some commentators were under the impression that we dismiss traits or that we equate traits to a single item (e.g., **Asendorpf**). We do not. **Steyer** is absolutely right when postulating that trait theory and the network perspective are not incompatible. That is, if one is willing to let go of the idea that a personality trait necessarily is a *latent cause* of thoughts, feelings and behaviors, then traits and the network perspective are perfectly fit for marriage. We may have added to the confusion by discussing individual personality components in which we endowed them with state-like or trait-like properties. Importantly, we do not hypothesize that single components/items are traits; we did want to argue that personality components can be stable (i.e., trait-like) as well as being subject to change (i.e., state-like). In that respect, there is indeed some overlap between the network perspective and the latent state-trait model, as was noticed by several commentators (**Costantini & Perugini; Rothmund, Baumert & Schmitt; Steyer**). However, contrary to the network perspective, the latent state-trait model is a *temperature* model in which individual items are caused by a latent state variable, which is, in turn, partially caused by a latent trait variable. In this respect, it is useful to distinguish between factor analysis as a pragmatic tool to organize data and the results of factor analysis as a model to explain data. Although it is unlikely that factor analysis of personality items would result in a credible explanatory model, this does not imply that it cannot be a useful statistical tool. As such, although we think the explanatory power of the latent state-trait model is limited, we acknowledge its usefulness as a statistical tool that can help in determining which personality components exhibit more state-like properties compared with others.

In the network perspective on personality, there is ample space for traits such as extraversion and neuroticism, if these are interpreted as clusters of mutually reinforcing components. The main difference with current trait theories is that from this perspective, traits do not function analogously to temperature nor do the items function analogous to thermometers. Instead, we postulate that the constellations of components that we designate as signs of underlying traits in fact result from the direct, local interactions between personality components. These may or may not be equated to single items (a subject we will return to in the final sections of this chapter). We used the flocking behavior of birds as an analogy, which needs, given some of the comments, some additional clarification. It is important to stress here that we did not invent the idea that birds, and other species, display flocking behavior because of local rules. Many simulation studies have confirmed that from a set of simple rules—for example, steer towards average heading of neighboring birds—a complex flocking pattern (e.g., a V-shape) can occur (e.g., C. Hartman & Benes, 2006). Thus, there is no underlying flocking instinct (**Terracciano & McCrae**) or a latent 'seasonal change' variable (**Schimmack & Gere**) that explains the flocking behavior of birds. **Schimmack and Gere** are right when they stress, in defense of the latent trait view, that flocking behavior is not "...an independent entity that exists apart from the individual birds..." but that is exactly the point: in our view, that applies to personality traits as well. That is, extraversion is not an independent entity that exists apart from the individual extraversion components: instead, just as flocking,

personality traits *emerge* out of the interaction between personality components. As such, from a network perspective, the relationship between a trait and its components is not one of measurement but one of *mereology*: that is, extraversion components do not measure extraversion; the interactions between these components are what constitute extraversion.

What does this mean? For one, extraversion and other personality traits cannot be understood by meticulously studying the inner workings of a single personality component. We think that virtually all commentators would agree with this: we cannot, for example, understand neuroticism by discovering all there is to know about a single component such as feeling jittery. Yet that is exactly what the latent trait, temperature, model implies: most of what we know about temperature (what it is and how it is related to thermometers) can be discovered by precisely investigating how it is related to one particular thermometer (a mercury thermometer for example).

The pragmatic, biological placeholder: Defending the latent trait view

In defense of the latent trait view, some commentators deny the reification of latent traits. That is, they adhere to a factor model-based temperature view of personality traits but claim to refrain from endowing the latent variables with any realist connotation. When having a temperature model in mind and when philosophizing about the nature of personality traits, is it unavoidable to reify the latent variable (see also **Wilt, Condon, Brown-Riddell & Revelle**)? In principle, no. From a *pragmatic* point of view, it is possible—as **Lee** points out in the case of mathematical ability—to work with latent variable model without believing that the latent factor has a material referent. However, the moment one searches for biological determinants that correlate with the latent variable, or for heritability of the latent variable, one wades into the murky waters of reifying the latent variable at least to some degree. Although not explicitly—the majority of commentators would likely refrain from endorsing the statement that neuroticism resides in a particular structure in the brains of individual people—many personality psychologists implicitly reify the latent variable when claiming that neuroticism is highly heritable or that gene X is associated with being extraverted. For what would be the use of searching for genetic determinants of something one does not believe to exist in some shape or form? One cannot pinpoint the *location* of temperature either, yet climatologists who claim that a permanent increase in average temperature is associated with an upslope tree line shift do believe that temperature is a real and causal phenomenon, although they cannot directly observe or touch it. As such, although we agree with **Steyer** that strong reification of the latent variable of the sort that personality traits are believed to be in the minds of individual people might not be what the vast majority of personality psychologists think *when asked* (although Allport and his followers do commit to the hypothesis that traits are real, that they exist *in our skins*; Allport, 1968, p. 49; Funder, 1991), when correlating latent variables (by their sum score proxy) with all sorts of (non-)biological phenomena and by engaging in statements such as 'women are more extraverted than men', however; they do grant the latent variable a status that comes undeniably close to reification (see also Kievit et al., 2011).

In fact, that quest for biological and/or genetic mechanisms is often fueled by the desire to endow the latent variable with some realist connotation. In this vein, **Terracciano and McCrae** defend the latent trait view with an argument along the following lines: personality traits are heritable, they are thus biologically based mechanisms; and because they are biologically based, personality traits exist. First, it is misguided to

use heritability as evidence for the hypothesis that some aspect of human functioning is reducible to *specific* underlying biological processes. Turkheimer (1998) contrasts the silence of an ascetic monk and an aphasic individual as an example: both religiosity and aphasia are heritable traits, but everyone will agree that in the case of the monk, his/her silence, which is a 'symptom' of his/her religiosity, is not due to specific brain structures or processes (e.g., religious silence is caused by a lesion in Brodmann areas 10 and 24), whereas in the case of an individual with, say, Broca's aphasia, we know that his/her silence is caused by a lesion in Broca's area. Thus, the fact that neuroticism is heritable does not imply that neuroticism is reducible to/associated with specific biological mechanisms. Second, the more *general* statement that personality traits are biologically based mechanisms without implicating any specific structure or process is utterly uninformative. Ultimately, all behavior is biologically caused in some sense (i.e., the result of biological processes), and as such, biological reductionism of mental phenomena such as personality traits is pointless unless one would want to maintain the hypothesis that certain behaviors, thoughts and/or feelings are not ultimately grounded in the brain of the individual who experiences or displays them (see also Greenberg & Bailey, 1993; Kendler, 2005b).

For some commentators, it is not so much the supposed *biological reality* of personality traits that prompts them to defend the latent trait view. Rather, in what we call the *placeholder* argument, personality traits cannot be something other than latent variables because that is the only way to understand why certain behaviors/thoughts/feelings (1) are present in some but not all humans (**Terracciano & McCrae**); and (2) that do not appear to be causally related but are correlated (**Guillaume-Hanes et al.**). In this view, the latent variable functions as a placeholder for everything we do not (yet) understand (i.e., latent variable as an *unknown* phenomenon), which is notably different from the interpretation of the latent variable as it figures in measurement and structural models (viz, as an *unobserved* phenomenon). What is wrong with the placeholder argument? For example, **Terracciano and McCrae** argue that because some birds display flocking behavior whereas others do not; it *must* be so that an underlying *flocking instinct* exists that causes these behavioral differences between bird species. Let us translate this hypothesis into an example that pertains to humans: some women prefer high heels whereas others do not; thus, it must be so that an underlying *instinct to wear high heels* exists that causes these behavioral differences between women. This obviously makes no sense. Although it may well be that we do not (fully) understand why it is that some women prefer high heels whereas others do not, the reasons we can think of do not justify the need for an underlying instinct: high heels are not practical in certain jobs, some women wear high heels to look taller, high heels cause back problems in some women, etc. Naturally, there are examples where the latent placeholder would be more defensible, but the thesis that behavioral differences *necessitate* the existence of an underlying instinct/tendency is, in our opinion, highly questionable.

Now, suppose we would find a positive correlation between wearing high heels and working on the top floor of a skyscraper. According to **Guillaume-Hanes and colleagues**, this correlation can only be understood by introducing an underlying tendency, in this example something such as *elevation tendency*, because there is no sensible way in which one can justify a direct relation between the two behaviors. The latter part of this argument is true. Likewise, in their own example, it is virtually impossible that ice cream eating causes children to seek their teacher's approval as well as the other way around. Besides methodological reasons why ordinary correlations do not necessarily imply a true relation between two variables (e.g., large sample size that causes low correlations to become significant, partial correlation might reveal that correlation is caused by a third (non-latent) party, etc.), **Guillaume-Hanes et al.** ignore another reason why

two seemingly wildly removed phenomena are correlated, which does not involve latent entities: for example, most women who work on the top floor of skyscrapers take the elevator. And because they take the elevator, they are not bothered by the discomfort of wearing high heels when climbing stairs. As a result, these women more readily wear high heels than women who take the stairs to reach their lower floor offices or the other way around: some women in highly successful companies with predominantly male employees like to accentuate their femininity by wearing high heels. And successful companies often occupy the most expensive floors in skyscrapers, the top floors. As such, at the inter-individual level, two behaviors can be related through a *causal chain* that involves directly observable, non-latent variables.

Not as straight as an arrow: The real trouble for the latent trait view

Interestingly, the potentially most compelling argument in defense of the latent trait view was not once articulated by any of the commentators. That argument would have been that it is known how latent traits influence behaviors, thoughts and feelings; that is, that we know what the *arrows* in the measurement model signify. Consider again the analogy with temperature: we know exactly what the arrow between temperature and a measurement with a mercury thermometer means, namely, that an increase in ambient temperature results in an increase in the temperature of the mercury, causing (in a linear fashion) the mercury in the glass to expand. For personality traits, however, it is no surprise that this argument was not articulated because no one really knows how, say, neuroticism causes 'feeling jittery' and 'worries easily'. As Mischel and Shoda (1994) stated, if traits generate distinctive behavior, then evidence for this claim needs "... to be stated explicitly and announced clearly". Yet, to the best of our knowledge, this evidence is not unannounced, it simply is not there. Naturally, there are theories of how traits and behavior are linked, for example, the trait theory as postulated in McCrae and Costa (1995). However, the meaning of the arrows in their model is shrouded in mystery: they are not really discusses nor empirically verified and are endowed in the model with the vague label 'dynamical processes'. So, when **Rothmund and colleagues** state that "...without theories about the nature of these causal links among components, it seems premature to refuse the classical trait models", they forget that for these very classical trait models, not one empirically verified theoretical model about the causal links between traits and concrete behaviors exists.

Thus, as Pervin (1994) rightly pointed out, personality traits are regarded as explanatory concepts, yet "...explanations are not offered in other than trait terms", resulting in circular arguments such as extraversion causes party-going; John likes to go to parties because he is extraverted. Why is it that trait theorists have not searched for how personality traits exert their supposed causal powers onto lower level behaviors, thoughts and feelings? There are probably many reasons—beyond the scope of this chapter to discuss—but one reason might have to do with the manipulability of the supposed latent traits. An important way of investigating the explanatory causal power of a theory is to manipulate the hypothesized cause of a certain phenomenon X after which one assesses the impact of that manipulation on X. One problem with personality traits, besides the obvious ethical constraints on such a research design, is that in order to manipulate a trait, one has to have a fairly good idea of what a trait is. And, although regarded as a *human universal* (McCrae & Costa, 1997), we have already argued in earlier sections that in current trait theory, there is no validated hypothesis on the nature of personality traits. Also, trait theorists themselves strongly

argue in favor of the stability of these traits (Terracciano, Costa, & McCrae, 2006). That is, especially after age 30, personality traits are supposedly relatively stable and thus relatively insensitive to external manipulation (if possible at all). On a final note, when reviewing the literature, it also seems that trait theorists are generally not very interested in answering the question of how traits cause behavior. In fact, McCrae and Costa (1995), for example consider the causal link between traits and behavior as self-evident and, as such, seem to obviate any need for formulating and testing explicit hypotheses about this link:

"..., the causal argument is in principle clear: traits as underlying tendencies cause and thus explain (in general and in part) the consistent pattern of thoughts, feelings, and actions that one sees. This kind of argument is so consistent with philosophical construals of disposition [...], with the theories of psychologists from Allport to Eysenck, with the assumptions underlying classical psychometrics, and with common sense that it is hard to understand why it should be problematic."

In our view, future research into the network perspective on personality should commit itself to the systematic identification and analysis of causal links between personality components at both the inter-individual and intra-individual levels. In our target paper, as **Asendorpf** rightly notes, we have only scratched the surface of the many possible interaction between personality components and how one can go about in analyzing possible causal mechanisms. For some relations at the inter-individual level, the causal mechanism probably operates at a more psychological level (those examples were most frequently discussed in the previous chapter): in general, liking to meet new people causes some individuals to seek out events where new people can be met, and therefore, these individuals frequent parties. And in these cases, there is no need for reducing this mechanism to a more *biological* explanation (e.g., neuron group X firing in region A causes neuron group Y in region B to fire). In other cases, biological mechanisms will be more important (e.g., in psychopathology and how not sleeping causes fatigue). And in these cases, we welcome the suggestions of **Read and Miller** and of **Wilt and colleagues** for how to incorporate biological mechanisms into the network model, for example, by positioning such mechanisms between genes and personality components.

Between-subject generalizations do not necessarily correspond to causal mechanisms that characterize within-person functioning. As **Furr and colleagues** show in the case of borderline personality disorder, causal mechanisms for developing the disorder might be very different for two people with the same diagnosis. Pertaining to normal personality, it might well be, for example, that in general, party-going behavior is predominantly caused by liking to meet new people, but that John likes to go to parties because he wants to be in the centre of attention and that Chris frequents parties because he wants to raise money for his next film (**Asendorpf**). So when **Asendorpf** criticizes the network perspective by arguing that people might vary in the causes of their party-going behavior, he inadvertently mentions a phenomenon (i.e., intra-individual differences in why certain behavior is present) that flows naturally from the basic premises of network theory. In network theory, the personality networks of John and Chris can be structured entirely different with the same end result: both men are extraverted. In John's case, his party-going behavior is caused by another personality component, wanting to be in the centre of attention; in Chris' network, his party-going behavior is caused by an external event (an upcoming film). Thus, there may be as many sources of *extraversion* as there are events that cause people to like parties or lead them to make friends easily. Importantly, this is not readily explainable by the latent trait, temperature, view. For in this view,

personality traits cause behavior in the same way in every individual, just as temperature causes a reading on a mercury thermometer in exactly the same way regardless of whether the temperature is measured in the Himalaya Mountains or in someone's backyard.

The pieces that make up the personality puzzle

So, if traits are not the *fundamental units of personality* (**Wilt and colleagues**) what are the basic pieces that make up the personality puzzle? We have suggested that personality components might fulfill that role: behaviors, thoughts and feelings that are associated with a unique causal system. Naturally, and as we have stressed, the definition of personality components is a first step; there is ample opportunity for refining this definition. A good refinement would be to consider more than items from self-report questionnaires (e.g., **Guillaume-Hanes et al.; Rothmund et al.; Wilt et al.**). For example, **Denissen, Wood and Penke** suggest the inclusion of functionalist components such as the reward value of social situations. With that said, it can be debated whether such functionalist variables act as components in a personality system or, instead, function as external forces that push the personality system towards a certain attractor (e.g., as a moderator that influences strength between two personality components). Another potential refinement of our original definition is to consider a component as consisting of multiple items. For example, **Costantini and Perugini** suggest that personality components might be better defined at the level of facets (i.e., sub-traits one level the Big Five, for example assertiveness). We agree that in certain cases, multiple items might be part of the same component, and in that sense, a component might be a facet, and the resulting network might be considered to be a higher level sub-network. In that case, one pragmatically chooses to study relations between sub-networks without assuming them to be fundamental, just as one can study interactions between sub-systems in the brain without phrenological assumptions. However, grouping multiple items would only work when the assumption of local independence is warranted as in the case of multiple thermometers. For example, sleep problems might be assessed by asking the individual , asking his/her spouse and by administering a polysomnography. As such, a measurement temperature model applies, and in that case, a personality component might be a latent variable. This approach would also effectively deal with the problem of incorporating measurement error into the network model (**Asendorpf**).

Eat the pudding!

In the previous chapter and this one, we have articulated a network perspective on personality in which traits result from the mutual interactions between personality components. Additionally, particularly in this chapter, we have articulated many reasons why the currently dominant latent trait view is in trouble: traits are probably not latent entities nor do they appear to have explanatory power, rendering the status of the latent trait view as the grand unifying theory of personality problematic. Naturally, criticizing existing theories and suggesting new one necessarily generates critical responses. However, the *concern* that **Schimmack and Gere** expressed over our "...suggestion that network analysis provides an alternative account of classic personality constructs..." does not make sense in our view, for science cannot progress without regularly questioning the basic assumptions on which research traditions are founded. That is, in a healthy scientific field, no doors should be closed; indeed, alternative theories should be welcomed, formalized and tested adequately. Regardless of which theory will in the end paint the best picture of how human being develop unique and yet, in some way, similar

personalities, one should not prematurely throw away the pudding without eating from it first.

Chapter 8

A practical guide to network analysis

Abstract

In network approaches to psychopathology, disorders result from the causal interplay between symptoms (e.g., worry → insomnia → fatigue), possibly involving feedback loops (e.g., a person may engage in substance abuse to forget the problems that arose due to substance abuse). The present chapter examines methodologies suited to identify such symptom networks and discusses network analysis techniques that may be used to extract clinically and scientifically useful information from such networks (e.g., which symptom is most central in a person's network). We also show how network analysis techniques may be used to construct simulation models that mimic symptom dynamics. Network approaches naturally explain the limited success of traditional research strategies, which are typically based on the idea that symptoms are manifestations of some common underlying factor, while offering promising methodological alternatives. In addition, these techniques may offer possibilities to guide and evaluate therapeutic interventions.

Adapted from: Borsboom, D., & Cramer, A. O. J. (*in press*). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology.*

Constructing and analyzing psychopathology networks

In recent decades, the construction and analysis of complex networks, which has its roots in physics and mathematics (Erdös & Rényi, 1959; Ising, 1925), has become a thriving enterprise in many fields that deal with complex organizations of mutually interacting entities. The problem of finding a way to analyze such systems has culminated in a set of powerful empirical research methods, generically known as network analysis, that can be applied to many different domains (Barábasi, 2011). One of the first papers to generalize the idea of marrying mathematical descriptions of network structures to diverse sets of data-driven networks was the classic paper by Watts and Strogatz (1998), which led to an avalanche of empirical and mathematical research on the structure and dynamics of complex networks. A good introductory text on the resulting literature is Newman (2010), while Kolaczyk (2009) and Barrat and colleagues (2008) provide excellent treatments of the applications of network modeling in dynamic models. Grimmett (2010) provides a technical introduction to more complicated probabilistic models, while Boccaletti and colleagues (2006) give a reasonably comprehensive and readable treatment of network approaches in different fields.

Table 8.1: Legend for Figures 8.1, 8.3, 8.5 and 8.6

Abbreviation	Meaning	Belongs to disorder
depr	Depressed mood	MD
inte	Loss of interest	MD
weig	Weight problems	MD
moto	Psychomotor disturbances	MD
repr	Self-reproach	MD
suic	Suicidal ideation	MD
mSle	Sleep problems	MD
mFat	Fatigue	MD
mCon	Concentration problems	MD
slee	Sleep problems	MD/GAD
conc	Concentration problems	MD/GAD
fati	Fatigue	MD/GAD
gSle	Sleep problems	GAD
gFat	Fatigue	GAD
gCon	Concentration problems	GAD
anxi	Chronic anxiety/worry	GAD
even	Anxiety about $>$ 1 event	GAD
ctrl	No control over anxiety	GAD
irri	Irritable	GAD
musc	Muscle tension	GAD
edge	Feeling on edge	GAD

At its core, a network is simply a set of elements (nodes) that are connected through a set of relations (edges). Nodes are usually visualized as circles and can represent any conceivable variable (e.g., symptoms, persons, airports, neurons). Edges connect these nodes and they can represent any conceivable sort of relationship (e.g., partial correlations, odds ratios, neuronal connectivity). To build a network, one first identifies the elements that will function as nodes. As an example, we use major depression (MD) and generalized anxiety disorder (GAD) symptoms (see Figure 8.1). Second, one determines what kind of relationship is represented by the edges. In panel A of Figure 8.1,

we define the relation as 'being part of the same disorder in DSM-IV': Any two symptoms that satisfy this relation are connected. These relations are coded in an *adjacency matrix* with all symptoms as rows (i) and columns (j). In this example, the matrix contains a 1 at position i,j if symptom i and j are connected, and a 0 otherwise. This matrix is subsequently used as input for visualizing the network. An alternative is shown in panel B of Figure 8.1 where the edges represent empirical correlations. In that case, the adjacency matrix equals the empirical correlation matrix. Thus, the construction and analysis of networks is highly accessible in the sense that the application of network models does not require extensive prior knowledge, as many other methodologies do: All one needs is a set of elements and an idea of how these elements are connected.

Table 8.2: Adjacency matrix pertaining to Figure 8.1

	depr	inte	weig	moto	repr	suic	slee	conc	fati	anxi	even	ctrl	irri	musc	edge
depr	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
inte	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
weig	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
moto	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
repr	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
suic	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
slee	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
conc	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
fati	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
anxi	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
even	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
ctrl	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
irri	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
musc	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
edge	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1

Psychopathology networks can be constructed in several ways, each of which may yield important information about the structure of disorders. For instance, as highlighted above, one can use the information diagnostic systems themselves, as these often contain clues about the causal constitution of disorders. Second, one can use the assessment of (causal) relations between symptoms, as rated by clinicians or patients. Third, one may use data on symptom endorsement frequencies to extract empirical patterns of association that can serve as input for network structures; for example, odds ratios, (partial) correlations, or pathways detected through causal search algorithms (Spirtes, Glymour, & Scheines, 2000). Below, we illustrate how such networks can be constructed and analyzed with existing data.

Networks based on diagnostic systems

Diagnostic systems such as the DSM-IV or the ICD-10 can be considered to partly reflect the structure of psychopathology through patterns of symptom overlap. A straightforward way of studying such patterns is by representing individual symptoms as nodes in a network and connecting then whenever they feature as symptoms of the same disorder (see panel A of Figure 8.1). This type of network reveals the structure of the diagnostic system itself. For instance, Borsboom and colleagues (Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011) used exactly the same procedure as was used for this figure to analyze the full symptom space of the DSM-IV. Figure 8.2 shows the resulting network, which represents patterns of symptom overlap in the DSM-IV. One

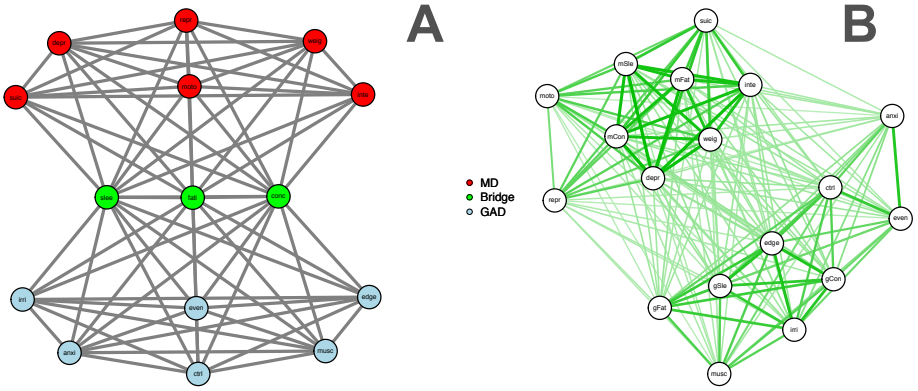

Figure 8.1: Networks for symptoms of major depression (MD) and generalized anxiety disorder (GAD) based on (A) the fourth edition of the DSM-IV and (B) correlations based on the National Comorbidity Survey Replication data. A: The symptoms of MD are placed at the top of the graph, bridge symptoms (i.e., symptoms that feature in both disorders) are in the middle, and GAD symptoms at the bottom. Symptoms are connected with a gray edge if they are part of the same disorder. Such a connection is coded in the adjacency matrix as a 1; no connection is coded as a 0 (see also Table 1). B: The edges represent correlations. The higher the correlation, the thicker the edge. The position of the nodes in the network is based on an algorithm, which causes strongly correlated symptoms to cluster in the middle, whereas symptoms with weaker connections to other symptoms figure more in the periphery of the figure (Fruchterman & Reingold, 1991).

striking feature of this network is the emergence of a *giant component*—a large group of nodes that are all connected to one another, either directly or via intermediary nodes (Newman, 2001b)—in which symptoms of mood (pink nodes), anxiety (orange nodes), and substance abuse disorders (green nodes) predominantly feature.

The giant component in Figure 8.2 has the characteristics of what is known as a small world in the network analysis literature (Watts & Strogatz, 1998); that is, on average, paths from one node to another are short and there is a large degree of clustering (i.e., the extent to which nodes tend to form a connected group). Most people are familiar with this idea through the work of Milgram (1967), who was among the first to demonstrate the small world phenomenon empirically. Milgram famously instructed people to send letters to other people (i.e., targets) they did not know by giving the letters to acquaintances they felt might know the target (or to somebody who might know somebody who...). The people who received the letters then did the same. On average, it took six steps to reach the target, a result that became famous as "six degrees of separation". Thus, a small world structure implies that, even though a network may be very large and feature strong clustering, any node can be reached from any other node in only a few steps. For the DSM-IV network, the small world property means that comorbidity appears to be, at least partially and in particular for mood, anxiety, and substance abuse disorders, encoded in the structure of the diagnostic criteria themselves (Borsboom, 2002).

How does one compute the two characteristics that determine whether a given binary network displays small world features, average path lengths and clustering? The shortest path length (SPL) between two nodes is the minimum number of edges that have to be traversed to reach one node from the other; for example, the SPL between chronic

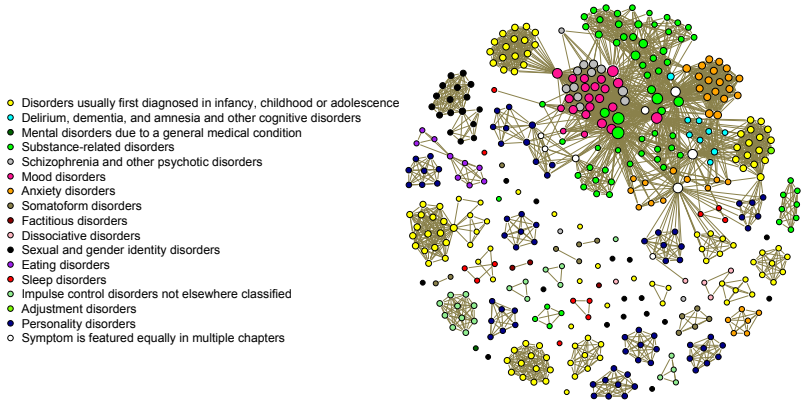

Figure 8.2: The DSM-IV symptom space. Symptoms are represented as nodes and connected by an edge whenever they figure in the same disorder. The color of the nodes represents the DSM-IV chapter in which they occur most often.

anxiety (*anxi*) and fatigue (*fati*) in panel A of Figure 8.1 equals 1, whereas it is 2 for the SPL between chronic anxiety (*anxi*) and depressed mood (*depr*) because they are not directly connected. The average shortest path length (ASPL) is the average of SPLs of all node-node pairs. In panel A of Figure 1 8.1, for instance, the ASPL is 1.253. Another measure of network size is the *diameter* of a network: the maximum path length between nodes in a graph. For the network in panel A of Figure 8.1, the diameter equals 2. The *clustering coefficient* C_i can be computed as follows. Suppose that a node i has k_i neighbors (the number of nodes with which node i is connected); then the maximum number of connections between these neighbors (MAX) equals $k_i(k_i-1)/2$: e.g., in panel A of Figure 8.1, irritable (*irri*) has eight neighbors, so MAX = 28 whereas sleep problems (*slee*) has 14 neighbors, so MAX = 91. C_i is the proportion of MAX that is actually present in the network. In panel A in Figure 8.1, C_i for irritable (*irri*) is 1 (28/28) and 0.604 (55/91) for sleep problems (*slee*).

Although the network in Figure 8.2 directly represents the DSM-IV rather than the structure of mental disorders, it is not entirely unreasonable to suspect that the network may harbor relevant causal information. This is because the DSM itself frequently mentions (or even requires) causal relations between symptoms of the same disorders. For instance, for the diagnosis of panic disorder (PD), it is required that a person has panic attacks, worries about the implications of these attacks, and changed his or her behavior as a result of panic attacks. In this case, the latter symptoms clearly depend causally on the presence of panic attacks themselves, so much so that this dependence is required for the diagnosis (i.e., worry about something distinct from panic attacks does not count as a symptom of PD). Similar constructions arise for the diagnosis substance use disorder (SUD), in which it is required that the person experiences problems as a result of substance abuse; post traumatic stress disorder, in which it is required that a person reexperiences traumatic events (in this case, the traumatic events are among the causes of reexperiencing them); and specific phobia, in which all other symptoms (e.g., enduring the phobic situation with intense anxiety) causally depend on the first symptom of being excessively fearful of a particular object or situation. In addition, causal links that are not explicated in the system may occasionally be highly likely. For instance,

in obsessive-compulsive disorder, compulsions are considered to be a means of reducing distress caused by obsessions (M. E. Franklin & Foa, 2011; e.g., a person washes his or her hands compulsively to reduce the distress caused by the obsession with cleanliness). Thus, at least for a subset of psychopathology symptoms, it is possible that their causal connection is in fact the reason that they figure as symptoms of the same disorder.

As we view it, diagnostic systems like the DSM are thus *not* theoretically neutral, as has been claimed to be the case from DSM-III onward (Maser, Kaelber, & Weise, 1991; Wakefield, 1997). Rather, this diagnostic system is replete with clinically relevant causal relations like the ones outlined above. At the level of causal relations, therefore, the DSM does theorize and, at times, it does so to a great extent when it comes to the causal order of symptom development. It is important to note that psychometric analyses of systems like the DSM-IV with latent variable models that simply ignore such clinically relevant causal relation, which are explicated in the system itself, should be viewed with caution.

Perceived causal relations

A second way of gaining insight into the causal organization of disorders is by asking experts of patients to report causal relations between symptoms. To our knowledge, the first researchers to ask experts about perceived causal relations between symptoms were Kim and Ahn (2002). For disorders such as anorexia nervosa, antisocial personality disorders, and major depression (MD), they asked clinicians to draw a line between two symptoms whenever they thought these two symptoms were somehow related. The clinicians were specifically told that such relations could mean anything (non)causal, from "co-occurs with" to "causes". Whenever clinicians drew a line between symptoms, they were asked to indicate the strength of this perceived relationship on a three-point scale. We asked 12 Dutch clinicians to do the same for MD, GAD, and Mania. The results pertaining to MD are shown in Figure 8.3 in which we have used an algorithm that positions strongly connected nodes in the middle of the graph and the more weakly connected nodes in the periphery of the graph (Fruchterman & Reingold, 1991): for example, according to the Dutch clinicians the symptom "depressed mood" is important in the disorder because it has strong connections with most of the other symptoms in the network. In contrast, the symptom "weight problems" is perceived to be less important since the clinicians do not think it is strongly related to any of the other symptoms of MD.

Recently, Frewen and colleagues (2012) have developed a systematic approach to the investigation of causal relations between symptoms by means of questionnaires that may be administered to clients. They call this methodology Perceived Causal Relations scaling. In this method, a person first indicated which of a set of symptoms is present. Second, each combination of presented symptoms (i, j) is combined in a question that assesses whether i caused j (reciprocal causal relations are typically allowed). In this way, one essentially builds a self-reported adjacency matrix for all symptom-symptom relations. The matrix defines a network that represents the cognitive representation of the causal structure of disorders. One could also see the network as a self-generated hypothesis on the network structure of a patient's disorder. The extent to which such hypotheses are in fact accurate is an important question for further research. If they are, then perceived causal relations scaling may offer a cheap and quick way to a rough assessment of psychopathology networks that could be used to construct informed treatment interventions.

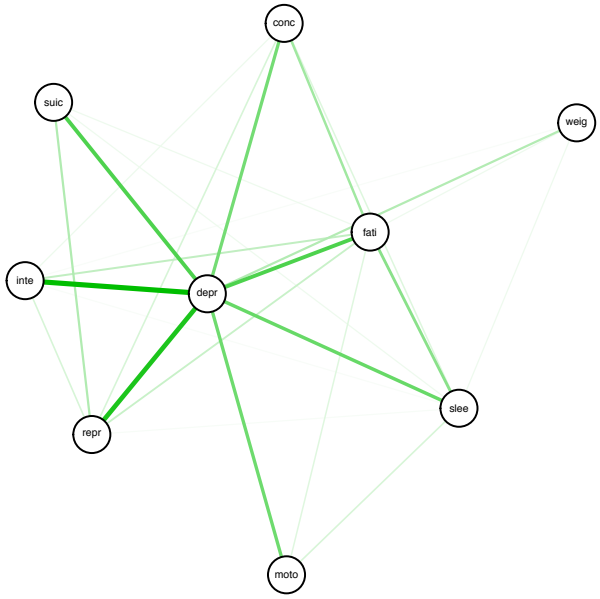

Figure 8.3: A network for MD based on the ratings of 12 Dutch clinicians. The nodes in the network represent the nine symptoms of MD; the edges between these nodes represent the mean connection strength between these symptoms as rated by the 12 Dutch clinicians (range: 0 = no connection; 3 = strong connection): The higher the mean rating, the thicker the edge. The position of the nodes in the network is based on an algorithm, which causes strongly correlated symptoms to cluster in the middle, whereas symptoms with weaker connections to other symptoms figure more in the periphery of the figure (Fruchterman & Reingold, 1991).

Extended psychopathology networks

Networks for psychopathology feature relations between symptoms. Typically, we see these symptoms as interacting with one another at the level of the individual person. However, in some cases, one person's symptom may 'infect' another person. Perhaps the most famous example of such a situation is the shared psychotic disorder or *folie deux*. This disorder may involve the development of a delusion in one person, who then infects another person through social communication. For instance, suppose that Bob becomes convinced that a government agency is spying on him. As a result of this symptom, Bob may keep the curtains closed, refuse to open the door, etc. Thus, Bob's primary symptom causes other symptoms, resulting in a network structure of psychopathology. Now imagine that Bob succeeds in convincing his spouse, Alice, of the veracity of his suspicions. As a result, Alice may also start withdrawing from social life and may develop symptoms similar to Bob's. Thus, the activation of Bob's symptoms not only has produced other symptoms within his own system, but has also produced symptoms in another person. We propose to call such symptom networks *extended psychopathology networks*.

Extended psychopathology networks may be studied in more or less the same way as ordinary psychopathology networks, but are especially useful when time information is present, so that one can estimate person-specific networks (see section "The many roads to disorder: individual networks) as well as the way they interact. Such methodology

could be used to chart the interaction between symptoms of different people in various social situations. This would be relevant for childhood psychopathology, for instance, because it would allow one to study the interaction between parents and children as problems develop over time. To give one example, in the study of developmental psychopathology, reciprocal interactions may exist between sleep problems and behavioral problems (Patzold, Richdale, & Tonge, 1998). Sleep problems of a child invariably lead to sleep problems of the parents; in turn, prolonged periods of poor sleep and behavioral problems in a child may lead to parental stress (Hoffman et al., 2008), which may result in less adequate handling of the child at bedtime, and hence leading back to sleep problems. Thus, in this case we have a feedback cycle that runs over symptoms that belong to various members of a family, and we see that the problems of neighboring individuals become intertwined.

Extending this idea further, one quickly reaches the conclusion that, in almost any mental disorder, significant social effects of this kind exist; in general, prolonged severe problems lead to a greater degree of social isolation. This means that the way in which one person's symptom network interacts with other people's networks leads to the alteration of that person's social network. Thus, carrying the idea of extended networks a little further, one sees a Russian doll of networks that are nested within other networks. Note that, even at this level, reciprocal influence is likely to be the norm rather than the exception, for the development of social isolation due to a network of personal problems may itself induce a further burden, thereby further enhancing the very problems that caused the isolated state in the first place. Thus, the complexity of psychopathology not only involves complex reciprocal relations between symptoms but also between networks of symptoms and social networks.

In this view, one follows the ladder upward, from symptom networks to social networks. Naturally, one can also extend networks in a downward fashion. For instance, one may unpack MD into a network of symptoms such as depressed mood and sleep problems. However, if one unpacks the concept of a sleep problem itself, one concludes that the symptoms themselves are complexly structured, with feedback cycles between hormones, external cues, and behaviors that give rise to the circadian rhythm. Thus, the reality of psychopathology involves a Russian doll of networks within networks in several layers of complexity. The exploration of such layered network structures is within reach given current data-gathering possibilities, and we think that the simultaneous analysis of social, symptom, and physiological networks is one of the main research challenges for the near future.

Association and concentration networks

Another way of exploring the causal organization of mental disorders is by studying empirical associations of symptom reports in patient or community samples. For instance, the matrix of correlations between symptoms is a symmetric symptom by symptom matrix, and as such, it can be treated as a weighted adjacency matrix. Panel B of Figure 8.1 shows such a network for symptoms of MD and GAD, in which the edges represent empirical correlations based on data of the National Comorbidity Survey Replication (NCS-R). We interpreted missing values that arose from the skip structure of the questionnaire as absent symptoms and replaced these by zeroes, which seems a reasonable course of action given the way the DSM-IV is set up. Naturally, other courses of action are possible, but these fall outside the scope of this review.

Such association networks are very useful for seeing at first glance which clusters of symptoms tend to be strongly connected or not. However, if one is interested in knowing which of these symptoms are truly related (i.e., discovering the causal skeleton that gives

rise to a particular correlational structure), then correlations may not provide optimal information. That is because a high correlation between any two symptoms might be the result of (1) a true direct (possibly reciprocal) relation between these two symptoms, or (2) a third variable that causes both symptoms, or (3) selection on a common effect of the symptoms (Pearl, 2000). An example of the first possibility is a high correlation between decreased appetite and losing weight: Not only are we quite sure that a direct relation exists between these symptoms of MD, we can also be confident about the directionality of this relationship: decreased appetite → losing weight. This and other direct causal relations between symptoms (e.g., insomnia → fatigue; self-reproach → suicidal ideation) are likely to form the causal skeleton of MD. On the other hand, in the second case, for instance, we might find a high correlation between avoiding a phobic situation/object and feeling distress over having a specific phobia. Then it is possible (and perhaps likely) that these two symptoms are not directly related (neither avoidance → distress nor distress → avoidance) but that their association is caused by a third symptom of specific phobia: exposure to the phobic situation/object provokes intense fear, as a result of which a patient (1) avoids the phobic situation/object and (2) feels distressed about the whole situation (thus exposure → avoidance and exposure → distress). As such, in this example, a direct relation between avoidance and distress might not be part of the causal skeleton of specific phobia. Compare this situation with smoking: Having yellow-stained fingers and a nasty cough are—when sampled in a normal population—probably highly correlated but not because they are directly related. Their association instead arises because they are caused by the same phenomenon, namely, smoking.

How can one figure out which correlations are indicative of direct causal relations and which are not? In a first step, one may obtain the matrix of partial correlations—that is, the correlations between pairs of symptoms that remain when all other symptoms are controlled for—which may be considered to provide clues about the causal skeleton of a network (an undirected pattern of direct relations between variables). For example, one computes the correlation between X (e.g., yellow-stained fingers) and Y (e.g., having a nasty cough) given Z (e.g., smoking): If the resulting correlation approaches zero, then one has a good indication that X and Y are not directly related. Figure 8.4 shows such a partial correlation network (only partial correlations > 0.10 are depicted as edges in the figure) for the five symptoms of specific phobia, in which each correlation was computed when all other variables in the network are controlled for.

One sees, for example, that no substantial partial correlations remain between avoidance (*avoi*) and distress (*dist*), whereas a rather large correlation remains between exposure (*expo*) and avoidance (*avoi*). Such a partial correlation network is called a concentration graph (Cox & Wermuth, 1993). Note that partial correlations between dichotomous variables are not statistically optimal and should be interpreted with some care; on the other hand, in our experience, more elaborate statistical methods tend to paint a qualitatively similar structure —just like Pearson correlations between dichotomous variables (point biserials) lead to roughly similar structures as more elaborate coefficients such as log odds ratios or tetrachoric correlations. To the extent that this generalizes, network structures may be reasonably recovered from such approximations even though point estimates and standard errors for the relevant association coefficients may be inaccurate. Further methodological investigations are needed to determine to what extent this is true. Note that it is straightforward to lift this limitation by applying nonparametric conditional independence tests.

In the traditional disease model (i.e., common cause model), the most interesting individual differences are to be found at the level of risk factors/dysfunctions that cause a particular disease, although, naturally, individual differences also exist at the level of the symptoms. That is, for instance, cancer research is dedicated to elucidate (1) which

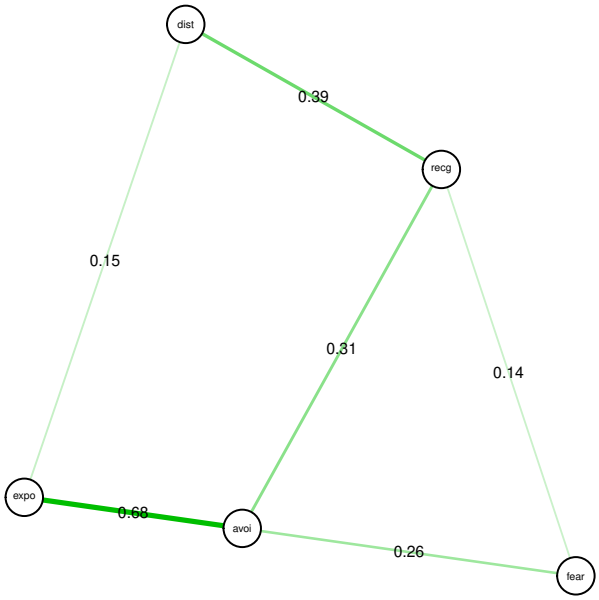

Figure 8.4: A network for specific phobia based on NCS-R data. The nodes in the network represent the symptoms of specific phobia; the edges between these nodes represent partial correlations > 0.10: The thicker the edge, the higher the partial correlation. The value of each partial correlation is placed on top of its corresponding edge. The position of the nodes in the network is based on an algorithm that causes strongly correlated symptoms to cluster in the middle while symptoms with weaker connections to other symptoms figure more in the periphery of the figure (Fruchterman & Reingold, 1991). *avoi*: the phobic situation is avoided or endured with intense anxiety or distress; *dist*: marked distress about having the phobia or avoidance/anxious anticipation/distress in the feared situation interfered significantly with the person's life; *expo*: exposure to the feared situation almost invariably provokes anxiety, which may take the form of a situationally bound or predisposed panic attack; *fear*: marked and persistent fear that is excessive and unreasonable, cued by the presence or anticipation of a specific object or situation; *recg*: the person recognizes that this fear is excessive or unreasonable.

risk factors predispose someone for developing, say, lung cancer (e.g., smoking, working with asbestos) and (2) why when two people smoke, one does develop lung cancer and the other does not (e.g., a genetic mutation). Such research is not aimed at elucidating why one person with lung cancer does complain of chest pains (i.e., a symptom) while another patient with the same disease does not.

From a network perspective, these assumptions about individual differences change radically, because a network perspective predicts that relevant differences arise at the level of the symptoms and the relations between them rather than at the level of the disorder. Concentration graphs in particular are useful for an assessment of which pathways between symptoms appear to be common. That is, strong partial correlations in a between-subjects weighted network (like the one in Figure 8.4 may indicate that these pathways reflect real causal relations that are relatively common in the sample on which the network representation is based. For MD, for instance, one may find that a common trajectory runs via depressed mood, loss of interest, and fatigue (see Figure 8.3). This formulation of common trajectories in terms of symptoms and relations between them

deviates markedly from existing perspectives on pathways to disorder. Naturally, such identified pathways would need to be validated in another independent sample.

In order to analyze such concentration (and association) graphs, the computation of path lengths and clustering coefficients needs to be generalized for weighted adjacency matrices. For computing shortest path lengths, Newman's (2001a) and Brandes' (2001) implementation of Dijkstra's (1959) algorithm minimizes the inverse of the distance between nodes i and j, formally defined as:

$$d^w(i,j) = \min(1/w_{ih} + ... + 1/w_{hj}) \tag{8.1}$$

$d^w(i,j)$ = distance between nodes i and j
w_{ih} = weight of the edge between nodes i and h

This makes sense when the weights are, for instance, correlations: the higher the correlation (and thus the stronger the hypothesized connection) between two symptoms, the smaller $d^w(i,j)$ and thus the shorter the distance. In Figure 8.4, for example, the shortest path from *fear* to *recg* is *not* the direct path ($1/0.14 = 7.142$) but rather the path via *avoi*: $1/0.26 + 1/0.31 = 7.071$. In this definition of a shortest path, only the weights are considered, not the number of edges. However, in some cases, it makes sense to include both the weights as well as the number of edges in computing shortest distances: for example, if one would be interested in the *fastest* and not so much the shortest path from MD to GAD in a comorbidity network, then it might be useful to take the number of edges (as a crude measure of speed) into account as well; so that, for instance, a 'faster' path containing three edges with relatively small weights would be preferred over a 'slow' path containing five edges with relatively large weights. Therefore, Opsahl, Agneessens and Skvoretz (2010) generalized the above formula to obtain:

$$d^{w\alpha}(i,j) = \min(1/(w_{ih})^\alpha + ... + 1/(w_{hj})^\alpha), \tag{8.2}$$

in which α is a *tuning* parameter that determines to what extent the number and weights of the edges count in computing the shortest paths between nodes. When α is 0, all denominators become 1 and as such, only the number of edges is taken into account. When α is 1, all denominators become the inverse of the weights (analogous to Dijkstra's algorithm) and in that case, only the weights of the edges are considered. For $\alpha < 1$ a path with fewer edges but with small weights would be favored over a path with more edges but with large weights. For $\alpha > 1$ the weights are more important than the number of edges.

Directed networks

Association and concentration graphs provide clues about possible causal relations between variables, but they do not provide information about the direction of causal relations (if these relations are unidirectional in the first place). The use of directed causal networks in statistical analysis has seen great developments in the past decades, especially in the work of Pearl (2000) and Spirtes et al. (2000). Unidirectional causal relations between nodes are typically represented by arrows. Causal analysis is easiest when the pattern of causal relations among variables creates a directed acyclic graph. In such a graph, all connections between nodes are directed, and it not possible to visit any node more than once when traversing the edges along the direction of the arrows in the graph (this means that there are no feedback loops). Under a (strict) set of statistical assumptions, the causal network structure can be deduced from a set of observational data by exploiting the connection between causal relations and certain patterns of conditional independence (a more detailed explanation is provided in Chapter 5). Danks and

colleagues (2010) have suggested that such explorative approaches, like the one implemented in the R package *pcalg* could be profitably used to build causal psychopathology networks. Figure 8.5 provides a graphical representation of the result of applying the PC algorithm (Spirtes & Glymour, 1991) to the NCS-R MD and GAD data with *pcalg*. The resulting network accords well with the idea that the covariance between MD and GAD is mainly a result of the bridge symptoms they share (Cramer et al., 2010), as the PC algorithm does not detect any paths between MD and GAD symptoms that are not mediated through their common symptoms. Also, the MD network bears some resemblance to the clinicians' networks in Figure 8.3. For example, in both networks, depressed mood (node number 1 in Figure 8.5) only has outgoing arrows, suggesting that this symptom might come early on the road to developing MD by triggering the development of other symptoms (although some caution is in order here, since in the NCS-R data, people are interviewed about other MD symptoms only if either depressed mood or loss of interest is present). On an additional cautionary note, the resulting graphs from search algorithms such as *pcalg* cannot contain feedback loops (e.g., $X \rightarrow Y \rightarrow Z \rightarrow X$); and yet, it is exactly these feedback loops that we suspect have the potential to push a given network into a disordered state (e.g., a depressed state).

The many roads to disorder: Individual networks

Between-subjects psychopathology networks are useful in, for instance, investigating the general structure of psychiatric disorders as they can generate testable hypotheses about trajectories toward developing a psychiatric disorder that are shared by individuals. However, such patterns of individual differences yield little insight when it comes to the question of how and why individual people develop disorders; for example, why Bob developed an episode of MD while Susan developed a phobic fear of spiders. In order to generate statements about the initiation, maintenance, and treatment of disorders of individuals, one needs to study the networks of individuals.

From a network perspective, each individual may have his or her own network, which comes with specific vulnerabilities or risk factors. Figure 8.6 show two MD-GAD networks for two fictitious persons, Alice and Bob. The figure shows that Alice and Bob differ quite markedly in terms of how they can potentially develop MD and GAD. For example, in Bob's case, the strongest pathway from MD to GAD runs via weight problems (*weig*), fatigue (*fati*), and feeling on edge (*edge*); in Alice's network, the progression of MD to GAD runs via depressed mood (*depr*), thoughts of suicide (*suic*) and irritability (*irri*). Since these are mere hypothetical examples of the many ways in which people can develop an episode of MD, what kind of data would we need to construct and analyze these individual networks? And what can we infer from these networks in terms of individual risk of developing a certain disorder?

Time series, time series, and time series

When the aim of network analysis is to construct disorder networks for individuals, cross-sectional data will be of little use. That is, in the networks of individuals, an arrow between any two symptoms—say insomnia $\rightarrow$ fatigue—is indicative of a process that takes place over time (e.g., insomnia develops at time t whereas the fatigue is caused by this insomnia at a later point in time, say, at time $t+1$). As such, querying a person about his or her symptomatology at one point in time is simply not enough to extract the causal information necessary to build a network of this person's symptom space. As mentioned in the previous section, it is possible to ask people to draw their own causal scheme, but of course the success of such a method relies on the ability of people to accurately report

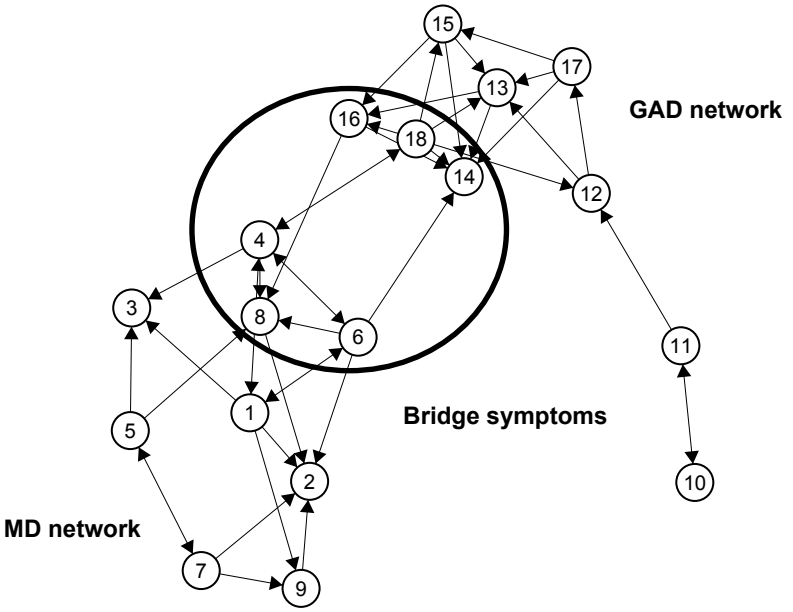

Figure 8.5: The directed MD-GAD network, based on the NCS-R data. Each edge represents a putative causal relation that remained after a search algorithm (pcalg) tracked all the possible conditional independence relations present in the data. If two symptoms are not directly connected, this implies that they are independent conditional on a subset of other symptoms. Double-headed arrows represent connections for which the algorithm cannot settle on a direction. *1*: depressed mood; *2*: loss of interest; *3*: weight problems; *4*: sleep problems (major depression); *5*: psychomotor problems; *6*: fatigue (major depression); *7*: self-reproach; *8*: concentration problems (major depression); *9*: suicidal ideation; *10*: chronic anxiety/worry; *11*: anxiety about more than one event; *12*: no control over anxiety; *13*: feeling on edge; *14*: fatigue (generalized anxiety disorder); *15*: irritable; *16*: concentration problems (generalized anxiety disorder); *17*: muscle tension; *18*: sleep problems (generalized anxiety disorder).

on their symptom development retrospectively, which may not be equally accurate in all circumstances (Henry, Moffitt, Caspi, Langley, & Silva, 1994).

A viable alternative is to collect time series data (Hamaker, Dolan, & Molenaar, 2005). That is, one asks individuals to report on various aspects of their physiological and psychological well-being at least once a day for many consecutive days. In one such recent research protocol, the experience sampling method (see Aan het Rot, Hogenelst, & Schoevers, 2012; Myin-Germeys et al., 2009, people are asked to report, during their normal daily life, their thoughts, feelings, and symptoms as well as the context in which these thoughts/feelings/symptoms take place and the appraisal of the context. One of the major advances of using such a method is that one is able to collect not only time-intensive data but also (1) data on the relation between events happening in a person's life and the subsequent ripple effects of that event in the symptomatology of this person, and (2) data from people without psychopathology who might be progressing toward developing a mental disorder. That is, in the latter case, one has excellent data to study why some people develop mental disorders while others do not, which in our opinion is the most pressing question in the entire realm of psychopathology.

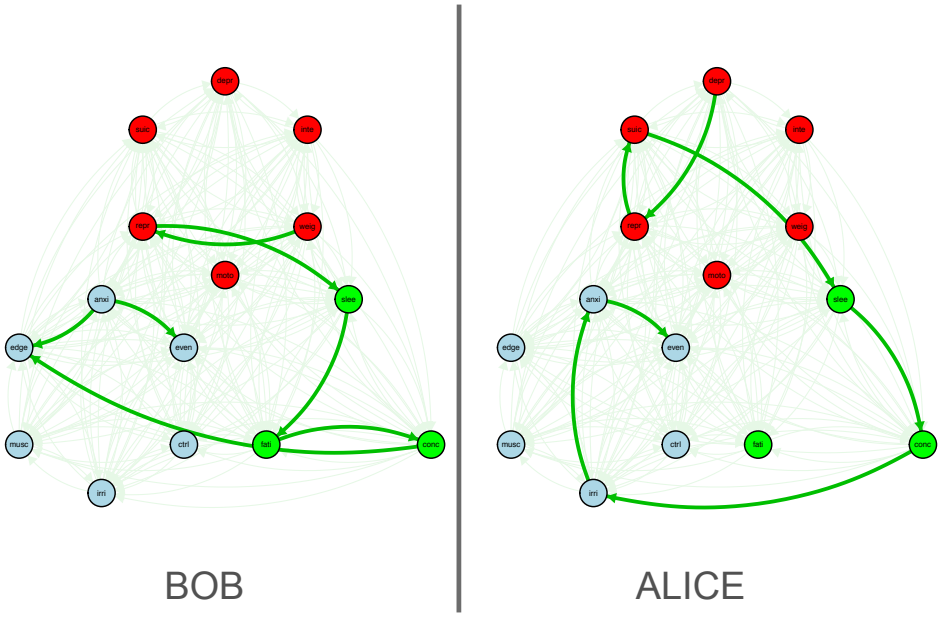

Figure 8.6: Hypothetical major depression (MD) and generalized anxiety disorder (GAD) networks for two fictitious people, Bob and Alice. Red nodes represent MD symptoms, green nodes represent bridge symptoms and blue nodes represent GAD symptoms. Thicker green edges represent stronger causal relations. These networks show that there are many ways to develop both MD and GAD.

Another possibility to learn about the intra-individual behavior displayed by a given network structure is by simulating time-intensive intra-individual data. With such simulated data, many of the interesting questions in psychopathology can be studied. For example, based on simulated data, Borsboom et al. (2011) showed that the percentage of "diagnoses" in simulated individual MD-GAD networks (in which comorbidity could only arise via bridge symptoms; see panel A of Figure 8.1) could account for prevalence rates for MD and GAD, comorbidity, and basic psychometric characteristics of the data at the same time. Also, based on a method to simulate data freely available online in the modeling environment NetLogo (Wilensky, 1999), one can study the known impact of stressors (e.g., negative life events such as the loss of a loved one) on individual symptoms of MD and relations between them (Cramer, Borsboom, et al., 2012; Keller et al., 2007). We developed such a simulation model of MD that can be found at http://ccl.northwestern.edu/netlogo/models/community/Symptom%20Spread%20Model (van Borkulo et al., 2011). In this model, virtually anything—from symptom development to stressors—can be manipulated by the user. It works as follows. At each time point, the model computes the probability of each symptom i to become activated at the next time point with the logistic function:

$$e^{\sum(ax-b)}/(1+e^{\sum(ax-b)}) \tag{8.3}$$

Here, $\sum(ax)$ is the activation sum of all symptoms at the previous time point (coded in a vector x) times the weight of the relevant connections (collected in a vector a) and could be seen as the total incoming effect for symptom i at that time point. b is a vector of

symptom-specific thresholds derived from the item difficulties of empirical data (Aggen, 2005). Two of the parameters that directly affect the probability functions in this basic setup and that can be altered by the user in real time are (1) number of connections (e.g., if all nine symptoms of MD are connected, then each symptom has eight neighbors) and (2) connection strength: The stronger the connections, the more influence the activation of symptoms has on other symptoms (thus directly affecting the a parameter vector in the model). Naturally, it must be noted that through such exercises one primarily learns something about what behavior is actually implied by one's theory, but in the case of network model this way of working can be quite revealing.

The analysis of time series

One can analyze time-intensive intra-individual data in a number of ways. The most straightforward way is to define connections in the network of an individual as representing the lag-1 correlations. That is, for example, if the network of Alice in Figure 8.6 would be a lag-1 correlation network based on empirical data, then the arrow from concentration problems (conc) to irritability (irri) means that concentration problems at time t predict irritability at time $t + 1$. Likewise, in Bob's network, the arrow from weight issues (weig) to self-reproach (repr) would mean that weight issues at time t predict self-reproach at time $t + 1$. If t would be measured in days, then a lag-1 correlation between feeling blue and eating more would probably be an appropriate time window; that is, it is plausible that feeling blue one day can make one eat more the next day. However, lag-1 correlations are probably not appropriate for other hypothesized relations between symptoms. For example, not sleeping for one night may not trigger fatigue immediately. Rather, one would expect a gradual build-up of sleepless nights, say, five, before true fatigue sets in. Thus, in this particular example, one would need to model this relationship in terms of a process that builds up over time.

Another option is to look at the entire available time window and define the connections between symptoms not in terms of (lag-1) correlations but rather in terms of the beta coefficients that result from a regression analysis through vector autoregressive modeling (Hamaker et al., 2007). For example, one could follow Bob for a prolonged period of time, assessing his depressive symptomatology every day on seven-point scales. Then, one could compute partial correlations in order to get a rough idea of the causal skeleton of Bob's network. With this information, one subsequently determines the neighbors of each symptom: All connections that represent a partial correlation of 0.10 or less (or some other optimum at which all nodes are connected with a minimum number of edges) are deleted. In Figure 8.4, for example, this procedure resulted in fear having two neighbors; that is, fear is connected with two other nodes in the network, not with all four. Next, one regresses each symptom at time t on its neighbors on t - 1 and calculates the regression weights. These weights would then represent the strength of the connections in Bob's network. In a next step, one could attempt to determine directed acyclic graph structures for this type of data (Eichler, 2007; Wild et al., 2007).

Thus, the analysis of time series could be executed in ways roughly similar to the previously discussed between-subjects data, but in this case to determine the network structure of the individual person. In general, a significant variety of models previously developed in econometrics and biometrics is available to construct network models (Kolaczyk, 2009). This may offer genuinely new ways of charting intra-individual network structures. Further developing methodology to do this in psychological applications would greatly facilitate research in this area. In addition, intra-individual network structures could offer novel ways of planning treatment, for instance by targeting the most important symptoms in a person's network structure.

Risk in individual networks

Regardless of how one defines the connections in the networks of individuals, what can we say about risk in terms of these networks? As mentioned previously, in disease, risk is defined at the level of the disease entity, which is not present in a network, at least not as an entity that is separable from its symptoms. From a network perspective, there are at least two ways in which a network can harbor risk of developing a certain mental disorder.

First, the structure of a particular network might be risky. To illustrate this concept, one may consider the symptoms of a disorder network to be domino tiles and view the connections between the symptoms as the distances between the domino tiles. Then relatively weak connections are analogous to domino tiles spaced rather widely apart (see left panel of Figure 8.7). As such, if, for instance, the symptom $X1$ in Figure 8.7 were to arise, then the probability of that symptom causing the development of other symptoms is relatively slim. In this case, the toppling of one domino tile will not likely result in the toppling of others, because they have relatively large distances between them. On the other hand, strong connections are analogous to domino tiles with short distances between them (see right panel of Figure 8.7). In that case, if symptom $X1$ were to be developed, then its activation would likely spread through the network like a virus spreads through a population: The toppling of that one domino tile will likely topple the other dominoes as well because of the short distances between tiles. Thus, suppose that Alex has had a drinking problem in the peat, which has caused all sorts of problems (e.g., financial problems, divorce from his wife), but at the moment he is sober after a successful intervention (see also Cramer et al., 2010). In that case, the structure of Alex's substance use network may still be risky, as there are only short distances between the domino tiles. Then, if Alex, for whatever reason, would have one drink, this would quickly culminate in other symptoms, such as financial consequences and problems with the important people in Alex's life. That is, a risky network structure might be the drive behind relapse, an important and well-known phenomenon in clinical psychology.

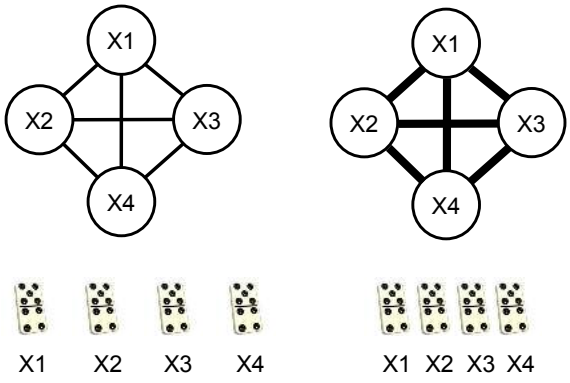

Figure 8.7: Two hypothetical networks of symptoms $X1 - X4$ (analogous to the domino tiles under the networks) that stand in certain relations toward one another (analogous to the distances between the domino tiles). In the left panel, the symptoms $X1 - X4$ are weakly connected, so in terms of the analogy, the distances between the domino tiles are relatively large. In the right panel, the symptoms $X1 - X4$ are strongly connected, so in terms of the analogy, the distances between the domino tiles are relatively short.

Second, there might be symptoms that, when developed in a particular person, have a stronger causal influence on the rest of the network compared to other symptoms. That

is, in reality, different symptom pairs will have different connection strengths, which determine the extent to which symptoms causally influence one another (as opposed to the networks in Figure 8.7, in which al connections within one network were equally strong). Consider, for example, the network depicted in Figure 8.8. In this network, $X1$ only has strong connections with the other symptoms in the network; that is, $X1$ is a central symptom in this network. On the other hand, the other symptoms, for example $X3$, have one strong connection but two weak connections; that is, $X3$ is a peripheral symptom (as are $X2$ and $X4$). Now, in terms of risk, the central nodes in someone's network are the most dangerous: If a central symptom becomes activated in someone, then the probability of that symptom causing the development of other symptoms is high (because the central symptoms are strongly connected to the other symptoms in the network); higher than when a peripheral symptom is activated.

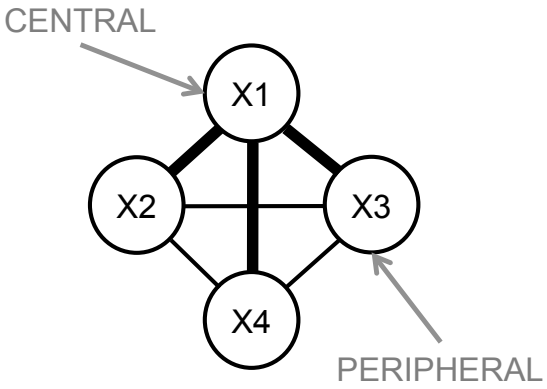

Figure 8.8: A hypothetical network of symptoms $X1 - X4$. Symptom $X1$ only has strong connections with the other symptoms in the network; this is a central symptom in this network. On the other hand, symptoms $X2 - X4$ all have one strong connection and two weak connections. These symptoms are thus more peripheral in the network.

There are multiple ways to compute the centrality of nodes in a network. One way is by computing the *closeness* (Opsahl et al., 2010) of node i, which is defined as the inverse of the total length of all SPLs between node i and all other nodes in the network; formally defined as:

$$C_C^{w\alpha}(i) = \left[\sum_{j=0}^{N} d^{w\alpha}(i,j)\right]^{-1} \tag{8.4}$$

where $d^{w\alpha}(i,j)$ is the shortest path length as explicated in equation 8.2;
N = total number of nodes in the network

For example, when one only considers the weights (i.e., $\alpha = 1$) in Figure 8.4, the closeness of avoidance (avoid) $= (\underbrace{1/0.68}_{\text{SPL[4,1]}} + \underbrace{1/0.26}_{\text{SPL[4,2]}} + \underbrace{1/0.31}_{\text{SPL[4,3]}} + \underbrace{1/0.31 + 1/0.40}_{\text{SPL[4,5]}})^{-1}$

where, for example, SPL[4,2] = shortest path from node 4 to node 2).

A downside of closeness centrality is that one cannot compute it when one or more nodes are not connected (e.g., in Figure 8.2, nodes that are not part of the giant component) because then the SPL between two nodes becomes infinitely large. Closeness

is in particular suited as a measure of centrality if one is predominantly interested in the capacity of, for example, a symptom to quickly activate a large number of other symptoms. Another measure of centrality is *betweenness* that quantifies how many a node i lies on the shortest paths between the other nodes in the network. Formally, betweenness including a tuning parameter α for node i is defined as follows (Opsahl et al., 2010):

$$C_B^w(i) = \sum \frac{g_{jk}^{w\alpha}(i)}{g_{jk}^{w\alpha}} \tag{8.5}$$

$g_{jk}^{w\alpha}$ = the number of shortest paths between two nodes[1];
$g_{jk}^{w\alpha}(i)$ = the number of those paths that go through node i

For example, in Figure 8.4, the betweenness of *fear* is 0, and 8 for *avoi.* This centrality measure might be particularly suited when assessing how one can go from one part of the network to another part, as might be interesting in the case of comorbidity (which nodes predominantly lie on the shortest paths between MD and GAD)?

Finally, another measure for computing the centrality of nodes is the *degree* or *node strength* which amounts to the the sum of the weights of the all edges incident in node i:

$$s_i = C_D^w(i) = \sum_j^N w_{ij} \tag{8.6}$$

In terms of risk of developing psychopathology, both this measure as well as closeness might be particularly suited as a means to determine which nodes in such a psychopathology network are central. When degrees are computed for each node, one can compute the relative frequencies of each degree in order to obtain a *degree distribution.* Most of empirical degree distributions will be highly skewed to the right, meaning that the majority of nodes in the network have a low degree, with a few exceptions (called *hubs*) of nodes with a high degree. Some networks are called *scale free* because their degree distribution approximately follows a power law: $P(k)\ k^{-\gamma}$ in which k is the degree and γ a particular constant. The World Wide Web, social networks and some biological networks appear to be real-world examples of scale free networks.

[1] If bidirectional then both path j-k and k-j

Chapter 9

Discussion

The reality of psychological constructs such as mental disorders and normal personality traits is not adequately captured by common cause models and their psychometric equivalent, latent variable models. In this dissertation, I have outlined a novel perspective, the network perspective, in which thoughts, feelings and behaviors directly interact with one another. In the case of mental disorders, the network approach is promising in that has the potential to explain many empirical phenomena such as comorbidity, the relation between stressful life events and individual symptoms of major depression, and spontaneous recovery. In the case of normal personality traits, the network approach offers an alternative theory of traits, in which they are a *consequence* of direct interactions between thoughts, feelings and behaviors. Additionally, the network perspective on normal personality has shown promise in linking genes to specific personality components as well as offering an alternative conceptualization of the relation between normal personality and psychopathology. While introducing the network perspective and outlining its viability for mental disorders and normal personality was the central aim of this dissertation, it is now time to take a step back and consider (1) (other) psychological constructs for which a network perspective might (not) be accurate and (2) some methodological obstacles that need to be overcome in order for firmly establishing network models as part of the psychometric toolbox.

Networks: Yes or no?

Future work will need to elucidate for exactly which disorders the network perspective is the best explanation and for which it is not. Throughout this dissertation, major depression was the prime example of a disorder for which we think a network model might be accurate in explaining its pathogenesis. Other examples include panic disorder. Panic disorder has four symptoms (APA, 1994): (1) recurrent unexpected panic attacks; (2) at least one of the attacks has been followed by one month (or more) of one (or more) of (2a) persistent concern about having additional attacks, or (2b) worry about the implications of the attack or its consequences (e.g., losing control, having a heart attack, "going crazy"); and (3) there is a significant change in behavior related to the attacks (e.g., avoiding public places: agoraphobia). Here, as explicated in Borsboom

Partly adapted from: Cramer, A. O. J., & Borsboom, D. (under review). Problems attract problems: The network perspective on mental disorders.

(2008b), a possible causal chain leading up to the pathogenesis of panic disorder is easily envisioned: 1 → 2a → 3 and 1 → 2b → 3. Another example of a disorder for which the network model might paint a realistic picture of its pathogenesis is bipolar disorder, especially the clinical manifestations in which patients switch between (hypo)mania and (sub-threshold) depression. For example, empirical research suggests that sleep disturbances are intimately associated with both the onset of a manic phase (i.e., decreased need for sleep is regularly observed before the mood switch to mania: e.g., Bauer et al., 2006; Leibenluft, Albert, Rosenthal, & Wehr, 1996; Plante & Winkelman, 2008) as well as with reducing depressive symptoms (i.e., sleep deprivation has antidepressant effects: e.g., Bunney & Bunney, 2012; J. Gillin, 1983; Hemmeter, Hemmeter-Spernal, & Krieg, 2010; Kuhs & Tölle, 1991; Landsness, Goldstein, Peterson, Tononi, & Benca, 2011; Wirz-Justice & van den Hoofdakker, 1999; Wu & Bunney, 1990). That is, a *symptom* (i.e. sleep disturbances) might be an important bridge between manic and depressive states, an observation that sits well with network models, as was shown in Chapter 2.

The origins of other disorders might not be predominantly founded upon symptom-symptom interactions. For example, in the case of development of affective symptoms after traumatic brain injury, it might be so that the symptoms are caused by the brain injury. That is, especially when the traumatic brain injury is located in the left anterior region, the injury itself might be the common cause of affective symptoms, which might culminate into an episode of major depression (Ownsworth & Oei, 1998). This is notably different from the situation in which affective symptoms are developed because of an inability to cope with the consequences of the traumatic brain injury (e.g., a person cannot adjust to the fact that the injury has resulted in the inability to return to work). In that case, the traumatic brain injury is a stressful life event for which we have evidence (see Chapter 4) that they can influence individual affective symptoms. In other disorders, it might be the case that symptom-symptom interactions are crucial in *maintaining* a pathological condition such as addiction (e.g., substance use → being broke → stealing from sister to buy substance → legal problems → substance use) but that the *initiation* of repeated substance use has its roots in an imbalance between dopaminergic circuits that underlie reward and conditioning and those that underlie executive functioning (Volkow, Wang, Fowler, Tomasi, & Telang, 2011).

Mental disorders, as we have portrayed them in our network models, are instances where an otherwise *normal* functioning system shows pathological behavior. That is, the nodes themselves are not pathological—not sleeping and feeling blue from time to time is normal and happens to virtually everyone—but for some ranges of parameter values (e.g., strong connections between symptoms) the behavior of the system becomes pathological. This is in line with how Mackey and Glass (1977) define a 'dynamical disease': "...characterized by the operation of a basically normal control system in a region of physiological parameters that produces pathological behavior". There are, however, probably disorders for which this definition of a disorder is inaccurate. Take, for example, psychopathy (e.g. Hare, 2003): psychopathy is a personality disorder that is characterized by symptoms such as pathological lying, manipulative, grandiose sense of self-worth and superficial charm. Now, for these types of symptoms, it is hard to argue that the symptoms are themselves normal and that it is the behavior of the system that is pathological in the case of psychopaths. To the contrary, in this case, it appears to make more sense to say that the symptoms themselves are pathological: most people are not manipulative and do not have a grandiose sense of self-worth. Another example is psychosis: a person who suffers from psychosis might have hallucinations (e.g., hearing the voice of your dead grandmother) or delusions (e.g., being convinced that your apartment is bugged by the CIA), both of which are symptoms that, as in the case of psychopathy, are themselves not normal: most healthy people will not have 'hallucinations' or 'delusions'

in their psychopathology networks (although the prevalence for hearing voices, for example, is higher—around 10%—than the prevalence of schizophrenia —around 1% van Os, Hanssen, Bijl, & Vollebergh, 2001). For these and similar disorders, the way we conceptualize and mathematically formalize corresponding network models might be inaccurate. Because the symptoms themselves are pathological, a network model for these disorders should, for instance, incorporate nodes that are continuous. For such nodes, we would subsequently have to show that, for example, simulation models in which the connection weights are equal across networks, activated nodes in the extreme range of the continuum (i.e., in the pathological range) are capable of producing pathological behavior of the system as a whole. On the other hand, one might question whether a network perspective, even if the weights are assumed to be equal for everyone, makes sense in the first place: if psychosis, psychopathy and similar disorders are pathological just because their respective symptoms are pathological and as a result, make it hard to function normally in our society, regardless of whether these symptoms interact with one another; then why the interactions? And without interactions, there is no network.

Methodological issues

One key issue for network theory to advance is the validation of techniques that are now in use to construct and analyze networks. We need to make sure, for example, that constructing networks in which the edges represent partial correlations and regression weights, are unaffected by, for example sample size and variance of the nodes. Getting this right is particularly important for future research endeavors aimed at elucidating differences in architecture between healthy people and people burdened with a particular mental disorder: what makes networks vulnerable to end up in a pathological state? When, for example, comparing the network of a healthy group with the network of a patient group (and the sample sizes of these groups are unequal), it is important to verify that differences between these networks are due to true differences in architecture instead of being the result of differences in, say, sample size: when defining an edge to be drawn whenever a (partial) correlation is significant, the network based on the larger sample size will contain more edges than the network based on the smaller sample size; but not necessarily because the architecture of these networks truly differs but because the larger the sample size, the more correlations are significant. Considering variance of the nodes, when for example comparing networks at multiple time points in which the edges are based on correlations, one could observe that correlations are increasing over time. This might be an indication that the connections between symptoms are truly becoming stronger. However, alternatively, it might also be an indication of restriction of range: when the variance of nodes is decreasing over time, this results in stronger correlations over time, which, in the case of restriction of range, have nothing to do with a true increase of connectivity between nodes.

Additionally, the adequacy of other techniques for determining the connectivity of a network should be investigated. For example, a common way of determining whether or not a causal relation exists between variables in econometrics is (non)linear Granger causality (Diks & Panchenko, 2005, 2006; Granger, 1969; Hiemstra & Jones, 1994). When considering a strictly stationary and weakly dependent bivariate time series process $\{(X_t, Y_t)\}, t \in \mathbb{Z}$ $\{X_t\}$ is said to Granger cause $\{Y_t\}$ if $F(Y_t|I_{t-1}) \nsim F(Y_t|I_{t-1} - X_{t-1})$ where F denotes a conditional probability distribution and I_{t-1} denotes a vector containing past values of X and Y (length of vector I depends on the length of the lag). Thus, Granger causality denotes the situation where the conditional probability distribution of Y_{t-1} given past values of X and Y is not equivalent to the conditional probability distribution of Y_{t-1} given only past values of Y: past values of X contain additional

information about current and future values of Y. In the linear multivariate Gaussian case, Granger causality can be relatively easily determined based on the significance of the coefficients of, for example, a vector autoregressive model. As such, when assuming linearity in the case that one has intensive time series data, one can use this technique to determine which nodes in a network might be causally related. In the non-linear (i.e., non-parametric) case, it is, at present, not possible to determine Granger causality for more than two time series.

Another important issue is the correspondence between inter- and intra-individual networks. While I have emphasized throughout the dissertation that collecting time-intensive data is the key to elucidating pathological mechanisms in individuals, the reality is that we still often deal with cross-sectional, inter-individual data. When estimating network parameters for such inter-individual data, is there anything one can conclude about the collection of intra-individual networks that underly the between-subjects network? For example, one might find in a large sample that the connection between depressed mood and thoughts of suicide is particularly strong. Does this finding reveal anything about the networks of the individuals that make up the sample: does it mean, for example, that, on average, most individuals in the sample also have a strong connection between depressed mood and thoughts of suicide in their network? It is probably too simple to, for instance, assume that individuals, in the binary case, who have a "1" for both symptoms necessarily have a strong connection between these symptoms. In future efforts, we will need to search for models with which we can estimate the likelihood of a response pattern (e.g., "1" for both depressed mood and thoughts of suicide) given a certain between-subjects network architecture. As such, we would be able to determine what intra-individual response patterns correspond to between-subjects connectivity (e.g., a "0" "1" response pattern is most likely for the connection between insomnia and fatigue with regression coefficient 1.2) and accordingly.

A final key issue concerns developing ways of quantifying the similarity of networks. For example, when one has constructed between-subjects networks for various time points during a clinical trial with an antidepressant, it might be interesting to determine whether the drug promotes actual change in the networks' architecture. When the connections are unweighted, the Jaccard index might be a likely candidate (Levandowsky & Winter, 1971), which is defined as the intersection of two sets, A and B, divided by their union: $J = \frac{(A \cap B)}{(A \cup B)}$. So for example, when both network A and B have 10 edges in total and they have five of them in common, $J = 5/15 = 1/3$. For weighted connections, J is not suited since, for example, J would be one if two networks are completely similar in terms of *which* connections are present, but this could obscure the fact that in network A, these connections are much stronger. A solution might be to weigh the union and intersection with the adjacency matrices of both networks (Gamallo, Gasperin, Agustini, & Lopes, 2001). In the case of two (or more) different and independent samples, for example when one wants to know whether the major depression architecture differs for men and women, other options exist. For example, one might construct a multi-group analysis in which the relative fit of a model in which network parameters (e.g., regression weights) are constrained to be equal across samples is compared to a model in which these parameters are estimated separately in both samples.

A final note

Humans are complex as are their thoughts, feelings and behaviors. In order to come to understand a small part of this complexity, it is necessary to construct models that reduce this complexity to such a level that it enables us to understand the essential pro-

cesses that result in certain behaviors without oversimplifying reality. As I have argued throughout this dissertation, current models of personality and of psychopathology are oversimplifying reality: while convenient in that such models are relatively easy to understand, they defy reality by designating traits and disorders, for the existence of which we have no evidence whatsoever, to be the ultimate glue that holds together behaviors such as liking people and liking parties; or symptoms such as hypersomnia and fatigue. If we really want to know why people with certain personality features struggle more in daily life and why some of these people develop psychopathology as a result, I do not think that current (trait) theories will prove crucial in finding the answers. A network perspective, with its emphasis on direct interactions between thoughts, feelings and behaviors that result in certain equilibria, might be helpful; especially when, for personality, we further develop our intuitions about what thoughts, feelings and behaviors are suitable components for personality networks.

Whichever theory of mental disorders one adheres to, they all share a deep desire to understand the inner workings of mental disorders. We all agree that finding out why some people are more vulnerable to developing mental disorders than others, how we can protect vulnerable people from harm, and how we can effectively treat people who have already fallen into the abyss of mental dysfunction are among the most pressing questions in the fields of clinical psychology and psychiatry. A disease model of mental disorders likely will not bring us any closer to finding answers to these questions. The network perspective, for some disorders, very well might.

Appendices

Appendix A

Abstracts commentaries in response to Chapter 2

Belzung, C., Billette de Villemeur, E., Lemoine, M., & Camus, V. (2010). Latent variables and the network perspective. ***Behavioral and Brain Sciences, 33*****, 150-151.**

We discuss the latent variables construct, particularly in regard to the following: that latent variables are considered as the sole explanatory factor of a disorder; that pragmatic concerns are ignored; and the relationship of these variables to biological markers is not addressed. Further, we comment on the relationship between bridge symptoms and causality, and discuss the proposal in relationship to other constructs (endophenotypes, connectionist-inspired networks).

Bornstein, R. F. (2010). The rocky road from Axis I to Axis II: Extending the network model of diagnostic comorbidity to personality pathology. ***Behavioral and Brain Sciences, 33*****, 151-152.**

Although the network model represents a promising new approach to conceptualizing comorbidity in psychiatric diagnosis, the model applies most directly to Axis I symptom disorders; the degree to which the model generalizes to Axis II disorders remains open to question. This commentary addresses that issue, discussing opportunities and challenges in applying the network model to DSM-diagnosed personality pathology.

Cervone, D. (2010). Aligning psychological assessment with psychological science. ***Behavioral and Brain Sciences, 33*****, 152-153.**

Network analysis is a promising step forward in efforts to align psychological assessment with explanatory theory in psychological science. The implications of Cramer et al.'s analysis are quite general. Network analysis may illuminate functional relations not only among observable behaviors that comprise psychological disorders, but among cognitive and affective processes that causally contribute to everyday experience and action.

Danks, D., Fancsali, S., Glymour, C., & Scheines, R. (2010). Comorbid science? ***Behavioral and Brain Sciences, 33*****,153-154.**

We agree with Cramer et al.'s goal of the discovery of causal relationships, but we argue that the authors' characterization of latent variable models (as deployed for such purposes) overlooks a wealth of extant possibilities. We provide a preliminary analysis of their data, using existing algorithms for causal inference and for the specification of latent variable models.

Davis, O. S. P., & Plomin (2010). Visualizing genetic similarity at the symptom level: The example of learning disabilities. ***Behavioral and Brain Sciences, 33*****, 155-156.**

Psychological traits and disorders are often interrelated through shared genetic influences. A combination of maximum-likelihood structural equation modeling and multidimensional scaling enables us to open a window onto the genetic architecture at the symptom level, rather than at the level of latent genetic factors. We illustrate this approach using a study of cognitive abilities involving over 5000 pairs of twins.

Fleeson, W., Furr, M., & Mayfield Arnold, E. (2010). An agenda for symptom-based research. ***Behavioral and Brain Sciences, 33*****, 157.**

The network approach proposed by Cramer et al. suggests fascinating new directions of research on mental disorders. Research is needed to find evidence for the causal power of symptoms, to examine symptoms thoroughly, to investigate individual differences in edge strength, to discover etiological processes for each symptom, and to determine whether and why symptoms cohere into distinct mental disorders.

Haig, B. D., & Vertue, F. M. (2010). Extending the network perspective on comorbidity. ***Behavioral and Brain Sciences, 33*****, 158.**

Cramer et al. make a good case for reconceptualizing comorbid psychopathologies in terms of complex network theory. We suggest the need for an extension of their network model to include reference to latent causes. We also draw attention to a neglected approach to theory appraisal that might be usefully incorporated into the methodology of network theory.

Haslam, N. (2010). Symptom networks and psychiatric categories. ***Behavioral and Brain Sciences, 33*****, 158-159.**

The network approach to psychiatric phenomena has the potential to clarify and enhance psychiatric diagnosis and classification. However, its generally well-justified anti-essentialism view psychiatric disorders as inevitably fuzzy and arbitrary, and overlooks the likelihood that the domain includes some latent categories. Network models misrepresent these categories, and fail to recognize that some comorbidity may represent valid co-occurrence of discrete conditions.

Hood, S. B., & Lovett, B. J. (2010). Network models of psychopathology and comorbidity: Philosophical and pragmatic considerations. ***Behavioral and Brain Sciences, 33*****, 159-160.**

Cramer et al.'s account of comorbidity comes with a substantive philosophical view concerning the nature of psychological disorders. Although the network account is responsive to problems with extant approaches, it faces several practical and conceptual

challenges of its own, especially in cases where the individual differences in network structures require the analysis of intra-individual time-series data.

Humphry, S. M., & McGrane, J. A. (2010). Is there a contradiction between the network and latent variable perspectives? ***Behavioral and Brain Sciences, 33*****, 160-161.**

First, we question whether Cramer et al.'s proposed network model can provide a viable scientific foundation for investigating comorbidity without invoking latent variables in some form. Second, the authors' claim that the network perspective is radically different from a latent variable perspective rests upon an undemonstrated premise. Without being demonstrated, we think the premise is potentially misleading.

Hyland, M. E. (2010). Network origins of anxiety and depression. ***Behavioral and Brain Sciences, 33*****, 161-162.**

Cramer et al. contrast two possible explanations for psychological symptoms: latent variables (i.e., specific cause) versus a network of causality between symptoms. There is a third explanation: The reason for comorbidity and the reported network structure of psychological symptoms is that the underlying biological cause is a psychoneuroimmunoendocrine information network which, when dysregulated, leads to several maladaptive psychological and somatic symptoms.

Johnson, W., & Penke, L. (2010). The network perspective will help, but is comorbidity the question? ***Behavioral and Brain Sciences, 33*****, 162-163.**

Latent variable modeling has revealed important conundrums in the DSM classification system. We agree that the network perspective has potential to inspire new insights and resolve some of these conundrums. We note, however, that alone it cannot really help us understand etiology. Etiology, not comorbidity, is the fundamental question.

Krueger, R. F., DeYoung, C. G., & Markon, K. E. (2010). Toward scientifically useful quantitative models of psychopathology: The importance of a comparative approach. ***Behavioral and Brain Sciences, 33*****, 163-164.**

Cramer et al. articulate a novel perspective on comorbidity. However, their network models must be compared with more parsimonious latent variable models before conclusions can be drawn about network models as plausible accounts of comorbidity. Latent variable models have proven generative in studying psychopathology and its external correlates, and we doubt network models will prove as useful for psychopathology research.

Markus, K. A. (2010). Questions about networks, measurement, and causation. ***Behavioral and Brain Sciences, 33*****, 164-165.**

Cramer et al. present a thoughtful application of network analysis to symptoms, but certain questions remain open. These questions involve the intended causal interpretation, the critique of latent variables, individual variation in causal networks, Borsboom's idea of networks as measurement models, and how well the data support the stability of the network results.

McFarland, D. J., & Malta, L. S. (2010). Symptoms as latent variables. ***Behavioral and Brain Sciences, 33*****, 165-166.**

In the target article, Cramer et al. suggest that diagnostic classification is improved by modeling the relationship between manifest variables (i.e., symptoms) rather than modeling unobservable latent variables (i.e., diagnostic categories such as Generalized Anxiety Disorder). This commentary discusses whether symptoms represent manifest or latent variables and the implications of this distinction for diagnosis and treatment.

Molenaar, P. C. M. (2010). Latent variable models are network models. ***Behavioral and Brain Sciences, 33*****, 166.**

Cramer et al. present an original and interesting network perspective on comorbidity and contrast this perspective with a more traditional interpretation of comorbidity in terms of latent variable theory. My commentary focuses on the relationship between the two perspectives; that is, it aims to qualify the presumed contrast between interpretations in terms of networks and latent variables.

Ross, D. (2010). Some mental disorders are based on networks, others on latent variables. ***Behavioral and Brain Sciences, 33*****, 166-167.**

Cramer et al. persuasively conceptualize major depressive disorder (MDD) and generalized anxiety disorder (GAD) as network disorders, rejecting latent variable accounts. But how does their radical picture generalize across the suite of mental and personality disorders? Addictions are Axis I disorders that may be better characterized by latent variables. Their comorbidity relationships could be captured by inserting them as nodes in a super-network of Axis I conditions.

Rothenberger, A., Banaschewski, T., Becker, A., & Roessner, V. (2010). Comorbidity: The case of developmental psychopathology. ***Behavioral and Brain Sciences, 33*****, 167-168.**

In developmental psychopathology, differentiating between the coexistence and the clinical entity of two problem areas is of utmost importance. So far, logistic regression analysis has already provided helpful answers, as shown in studies on comorbidity of tic disorders. While the concept of *bridging symptoms* may be investigated adequately by both logistic regression and the network approach, the former (latent variable) seems to be of advantage with regard to the problems of multiple comorbidities and development.

Rubinsten, O., & Henik, A. (2010). Comorbidity: Cognition and biology count! ***Behavioral and Brain Sciences, 33*****, 168-170.**

We agree with Cramer et al. that pure cases of behavioral disorders with no symptom overlaps are rare. However, we argue that disorders do exist and the network idea is limited and limiting. Networks of symptoms are observed mainly at behavioral levels. The core deficit is commonly at the cognitive or brain levels, and there the story is completely different.

Staniloiu, A., & Markowitsch, H. J. (2010). Looking at comorbidity through

the glasses of neuroscientific memory research: A brain-network perspective. ***Behavioral and Brain Sciences, 33*****, 170-171.**

As psychiatric illnesses have correlates in the brain, it is surprising that Cramer et al. make almost no reference to the brain's network character when proposing a network approach to comorbidity of psychiatric diseases. We illustrate how data from combined neuropsychological and functional and structural brain-imaging investigations could inform theoretical models about the role played by overlapping symptoms in the *etiology* of psychiatric comorbidity and the pathways from one disorder to another.

Tzur-Bitan, D., Meiran, N., & Shahar, G. (2010). The importance of modeling comorbidity using an intra-individual, time-series approach. ***Behavioral and Brain Sciences, 33*****, 172-173.**

We suggest that the network approach to comorbidity (Cramer et al.) is best examined by using longitudinal, multi-measurement, intra-individual data. Employment of time-series analysis to the examination of the generalized anxiety and major depressive disorder comorbidity enables a detailed appreciation of fluctuations and causal trajectories in terms of both symptoms and cognitive vulnerability.

van der Sluis, S., Kan, K. -J., & Dolan, C. V. (2010). Consequences of a network view for genetic association studies. ***Behavioral and Brain Sciences, 33*****, 173-174.**

Cramer et al.'s proposal to view mental disorders as the outcome of network dynamics among symptoms obviates the need to invoke latent traits to explain co-occurrence of symptoms and syndromes. This commentary considers the consequences of such a network view for genetic association studies.

van Geert, P. L. C., & Steenbeek, H. W. (2010). Networks as complex dynamical systems: Applications to clinical and developmental psychology and psychopathology. ***Behavioral and Brain Sciences, 33*****, 174-175.**

Cramer et al.'s article is an example of the fruitful application of complex dynamic systems theory. We extend their approach with examples from our own work on development and developmental psychopathology and address three issues: (1) the level of aggregation of the network, (2) the required research methodology, and (3) the clinical and educational application of dynamic network thinking.

Wass, S., & Karmiloff-Smith, A. (2010). The missing developmental dimension in the network perspective. ***Behavioral and Brain Sciences, 33*****, 175-176.**

We welcome network theory as a tool for modeling the multidimensional interactions that characterize disease. However, we feel that Cramer et al. have neglected one important aspect: how diseases change over developmental time. We discuss principles such as *fan in*, *fan out*, *bottlenecks*, and *common pathways*, argue that modeling these developmental aspects can be vital, particularly in deriving properly targeted treatments.

Yordanova, J., Kolev, V., Kirov, R., & Rothenberger, A. (2010). Comorbidity in the context of neural network properties. ***Behavioral and Brain Sciences, 33*****, 176-177.**

Cramer et al.'s network approach reconceptualizes mental comorbidity on the basis of symptom space originating from psychometric signatures. We argue that the advantages of this approach need to be regarded in the context of the multi-level functional organization of the neural substrate, ranging from neurogenetic to psychometric. Neuroelectric oscillations are proposed as a level-integrating principle.

Zachar, P. (2010). The abandonment of latent variables: Philosophical considerations. ***Behavioral and Brain Sciences, 33*****, 177-178.**

Cramer et al.'s critique of latent variables implicitly advocates a type of scientific anti-realism which can be extended to many dispositional constructs in scientific psychology. However, generalizing Cramer et al.'s network model in this way raises concerns about its applicability to psychopathology. The model could be improved by articulating why a given cluster of symptoms should be considered disordered.

Appendix B

Abstracts commentaries in response to Chapter 6

Asendorpf, J. B. (2012). What do the items and their associations refer to in a network approach to personality? ***European Journal of Personality, 26*****, 432-433.**

It is hard to judge the potential usefulness of a network approach to personality research because Cramer et al. (2012) mix up applications to one individual, inter-individual differences and intra-individual processes. From each perspective, the network units, their associations and causal interpretations of such associations have a completely different meaning, and it depends on the particular perspective, the level of aggregation and whether one wants to model measurement error whether latent variables have a place in network models in personality research.

Ashton, M. C., & Lee, K. (2012). On models of personality structure. ***European Journal of Personality, 26*****, 433-434.**

We suggest that the description by Cramer et al. (2012) of traditional models of personality structure does not perfectly reflect the models actually endorsed by researchers. Personality researchers assume that many variables will have considerable secondary loadings and that the major personality factors will not account for all of the covariation among those variables. A model that includes common factors provides a more parsimonious explanation of covariation among personality variables than does a model consisting of network links only.

Costantini, G., & Perugini, M. (2012). The definition of components and the use of formal indexes are key steps for a successful application of network analysis in personality psychology. ***European Journal of Personality, 26*****, 434-435.**

Conceiving personality as a network provides an interesting theoretical framework and a promising methodological perspective. The application of network analysis to personality psychology however is not straightforward, and some issues require careful consideration. We argue that the definition of components within networks cannot be limited to single items, and more work is needed to reflect the inherently hierarchical and indefinite nature of components. Additionally, we argue that formal empirical indexes must be clearly defined and consistently used to describe properties of personality

networks.

Denissen, J. J. A., Wood, D., & Penke, L. (2012). Passing to the functionalists instead of passing them by. *European Journal of Personality, 26*, 436-437.

The paper by Cramer and colleagues illustrates how a network approach can model personality systems without positing causal latent factors such as the Big Five. We applaud this effort but argue that nodes should be distinguished on more than quantitative grounds (e.g., displayed centrality or connectivity). To realistically model the affects, cognitions and behaviors that constitute real personalities, organizing constructs such as needs and comparators seems necessary. Incorporating them requires greater consideration of functionalist personality theories that link together environmental features and adaptive behavior in meaningful and stable ways.

Furr, R. M., Fleeson, W., Anderson, M., & Mayfield Arnold, E. (2012). On the contributions of a network approach to personality theory and research. *European Journal of Personality, 26*, 437-439.

Understanding personality structure and processes is one of the most fundamental goals in personality psychology. The network approach presented by Cramer et al. represents a useful path towards this goal, and we address two facets of their approach. First, we examine the possibility that it solves the problem of breadth, which has inhibited the integration of trait theory with social cognitive theory. Second, we evaluate the value and usability of their proposed method (qgraph), doing so by conducting idiographic analyses of the symptom structure of borderline personality disorder.

Guillaume-Hanes, E., Morse, P., & Funder, D. (2012). Network models in the organization of personality. *European Journal of Personality, 26*, 439-440.

The network perspective illustrates an important cautionary point concerning the interpretation of inter-item relationships. However, its complexity comes at a prices, including a possible lack of robustness and replicability, and difficulties in interpretation and achieving psychological insight. The most interesting and important manifestations of personality are diverse and consequential behaviors that are related because they really do reflect common underlying traits. Thus, the target article can serve as a reminder of the importance of ranging beyond self-report questionnaires to the much more difficult, expensive and important world of behavior.

Lee, J. J. (2012). Common factors and causal networks. *European Journal of Personality, 26*, 441-442.

The target article touches upon some of the most difficult and essential questions in personality psychology. Questioning the notion of a common factor as an as-yet-unobserved common cause of a behavior domain's exemplars, the authors propose using graphical representations to inspire hypotheses of more complex causal structures. I do not find the case for the de-emphasis of the common factor model to be compelling for those behavior domains (cognitive abilities) with which I am most familiar. It behoves all personality psychologists, however, to question the foundations of their favored tools.

Read, S. J., & Miller, L. C. (2012). Sources of constraint on network equilibrium. *European Journal of Personality, 26*, 442-443.

We agree with the authors' key point that the standard trait approach to personality does not provide a method for understanding the causal structure of personality. Furthermore, their new technique for visualizing structure shows promise. However, although genetic analyses are important, we think that they are, by themselves, inadequate as a source of information/constraint for understanding causal structure. Close attention must also be paid to the biological systems they they influence, the structure of social situation and the dynamics of the interactions among them. We outline one possible approach to these issues.

Rothmund, T., Baumert, A., & Schmitt, M. (2012). Can network models represent personality structure and processes better than trait models do? *European Journal of Personality, 26*, 444-445.

We argue that replacing the trait model with the network model proposed in the target article would be immature for three reasons. (i) If properly specified and grounded in substantive theories, the classic state-trait model provides a flexible framework for the description and explanation of person x situation transactions. (ii) Without additional substantive theories, the network model cannot guide the identification of personality components. (iii) Without assumptions about psychological processes that account for causal links among personality components, the concept of equilibrium has merely descriptive value and lack explanatory power.

Schimmack, U., & Gere, J. (2012). The utility of network analysis for personality psychology. *European Journal of Personality, 26*, 446-447.

We note that network analysis provides some new opportunities but also has some limitations: (i) network analysis relies on observed measures such as single items or scale scores; (ii) it is a descriptive method and, as such, cannot test causal hypotheses; and (iii) it does not test the influence of outside forces on the network, such as dispositional influences on behavior. We recommend structural equation modeling as a superior method that overcomes limitations of exploratory factor analysis and network analysis.

Steyer, R. (2012). Does network theory contradict trait theory? *European Journal of Personality, 26*, 447-448.

I argue that the trait and network theories of personality are not necessarily contradictory. If appropriately formalized, it may turn out that network theory incorporates traits as part of the theory. I object the opinion that if a trait is a cause of behavior, then it is necessarily an entity operating in the minds of individuals. Finally, I argue that liking parties can be a label for a random variable (item), a stochastic process (a family of items at different time points) and a latent variable (trait). In our colloquial language, we do not make these distinctions, which leads often to confusions.

Terracciano, A., & McCrae, R. R. (2012). Why do (some) birds flock? Causality and the structure of characteristic adaptations. *European Journal of Personality, 26*, 441-442.

Characteristic adaptations often cluster in mutually reinforcing networks. Evidence of stability and heritability suggests that the development of such networks is due

in part to the causal influence of enduring dispositions or traits. Many different genetic models are consistent with this hypothesis, and the quest for genes can be pursued at many levels; of these, the intermediate level of specific facets may be the most promising.

Appendix C

Where are the genes?

Abstract

This commentary was written in response to a target article by Johnson, Penke and Spinath (2011, published in the same issue of the European Journal of Personality) in which the missing heritability problem was discussed from many different angles. In this commentary, we present another reason for the apparent discrepancy between heritability estimates and gene-hunting results in psychopathological research (i.e., the missing heritability problem): if syndromes are networks of causally related symptoms in which both symptoms and relations between them are driven by different sets of genetic polymorphisms, then gene hunting based on a phenotypic sumscore might be ill-advised because it will only capture genetic variance that is shared among those symptoms and their relations.

Depressed parents predispose their children to become depressed as well. This phenomenon is not so much attributable to a depressogenic environment (inadvertently created by the parents) as it is due to the fact that major depression is a moderately heritable syndrome, with heritability estimates ranging between 37% and 60% (Boomsma et al., 2002; Kendler, Gatz, Gardner, & Pedersen, 2006; P. F. Sullivan, Neale, & Kendler, 2000). Combined with the high heritability of other mental disorders (Boomsma et al., 2002), it is surprising that despite many efforts, the genetic culprits have not been identified (see e.g., Sklar, 2002). For psychological traits in general, identified genetic polymorphisms typically account for less than 2% of the genetic variance (Levinson, 2006; Mitchell & Porteus, 2009).

The apparent discrepancy between high heritability and the inability to identify the responsible genetic polymorphisms has been termed the *missing heritability* problem and is pervasive in the realm of psychopathology (Manolio et al., 2009). In the same issue, Johnson et al. propose various plausible mechanisms that contribute to the missing heritability problem, ranging from methodological factors that might result in inflated heritability estimates to problems with the specific research strategies employed in gene hunting. Pertaining to the latter, in this commentary, we elaborate on a potential problem on which Johnson et al. did not reflect: what if the way we define a syndrome in current gene-hunting efforts is incorrect?

In psychopathological research, the most commonly used proxy for a phenotype is

Adapted from: Cramer, A. O. J., Kendler, K. S., & Borsboom, D. (2011). Where are the genes? The implications of a network perspective on gene hunting in psychopathology. *European Journal of Personality, 25*, 270-271.

the typical operationalization of a syndrome, that is, a sumscore (i.e., the total number of symptoms of a disorder present) that can be further dichotomized, using, for example, criteria as specified by the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV), to reflect the presence/absence of a particular disorder. In its most rudimentary from, genetic association studies identify genes or genetic variants as predisposing to a mental disorder if they predict the dependent variable in the design, that is, the (dichotomized) sumscore (van der Sluis, Kan, & Dolan, 2010). So far, this strategy has not been very successful at identifying the important genetic polymorphisms in the onset of mental disorders. In our view, this may be partly due to the fact that one of its most important assumptions—that a sumscore forms a valid representation of a syndrome—is fundamentally flawed.

Current approaches to gene hunting rely on the assumption that the relation between a phenotype, for example, major depression—and its observable attributes—for example, the A criterion in DSM-IV is one of *measurement*: a psychological phenomenon *causes* its observable attributes (e.g., extraversion causes party-going behavior: McCrae & Costa, 2008, p. 288). One of the far-reaching consequences of such a *common cause* view is that correlations among the observed attributes themselves are deemed spurious; they only exist because they share a common cause: mental rotation skills and verbal intelligence are only correlated because they share a common cause, namely general intelligence. In terms of searching for genes that are implicated in the onset of a syndrome, this view translates into the following chain of events: genes, via a host of hypothesized endophenotypes, result in a syndrome and that syndrome, because of a measurement relationship, causes a set of observable symptoms. But as we have argued elsewhere (Cramer et al., 2010), relations between symptoms might not only be non-spurious in nature but might also the very essence of what constitutes a syndrome (similar arguments have been made for general intelligence: van der Maas et al., 2006). For example, consider the correlation between two symptoms of major depression: insomnia and fatigue. Under the assumption of a common cause, the correlation between these two symptoms is spurious; it only arises because insomnia and fatigue share a common cause, major depression. It is, however, more likely to assume that this correlation exists because there is a real, straightforward causal relation between these two symptoms: if you do not sleep, you will become tired. Similar arguments can be made for a host of other psychological phenomena—for example, consider feeling comfortable around people and party-going behavior: do we need an overarching 'extraversion' trait to explain why these observed behaviors tend to covary?—and as such, it is premature to dismiss direct relations between observed attributes as being mere spurious by-products of an overarching construct. What does this mean for gene hunting efforts?

If constructs are indeed networks of causally related observables, individual differences are most likely to arise as differences in the strength of those relations: when Alice suffers from depressed mood, she fairly easily develops suicidal thoughts (i.e., strong relation between the observe symptoms 'depressed mood' and 'suicidal ideation'), whereas Bob does not ever contemplate suicide while feeling depressed (i.e., relatively weak relation). Furthermore, it is likely that the strength of such relations stands at least partly under genetic control. Now, it is not likely that each relation is influenced by the same set of genes for the sheer number of relations (k^2 - k in a network containing k observables/symptoms) in any given network greatly diminishes this possibility and the relations probably differ in terms of the endophenotypes (and thus genes) involved (e.g,. the more physiological homeostatic processes that are likely to govern relations between sleep van fatigue versus the more cognitive processes that are probably invoked in the relation between depressed mood and suicidal thoughts).

Hence, when trying to relate genetic polymorphisms to a sumscore, one only cap-

tures the genetic variance that is shared among those individual symptoms (including their relations); the different genetic polymorphisms that are responsible for individual differences in the strength of the *relations between* those symptoms are completely left unaccounted for. As such, the network approach may explain at least partly why current approaches cannot find the genetic culprits of mental disorders. By properly modeling their etiology, we increase our power to detect risk variants. It is, after all, the relations between symptoms that glue them together into a syndrome.

Appendix D

A constructionist account of emotional disorders

Abstract

This commentary was written in response to a target article by Lindquist, Wager, Kober, Bliss-Moreau and Feldman-Barrett (2012, published in the same issue of Behavioral and Brain Sciences) in which the brain basis of emotion was reviewed. Lindquist et al. present a strong case for a constructionist account of emotion. In this commentary, we first elaborate on the ramifications that a constructionist account of emotions might have for psychiatric disorders with emotional disturbances as core elements. Second, we reflect on similarities between Lindquist et al.'s model and recent attempts at formulating psychiatric disorders as networks of causally related symptoms.

Fear is not localized in the amygdala, nor does sadness exclusively arise in the anterior cingulate cortex. Unfortunately for Gall (Gall & Spurzheim, 1835), and more recent proponents, who hypothesized that single brain areas (later referred to as "particular circuits"; see Kandel & Squire, 1992) correspond to single functions (e.g., arithmetic skills), feelings (e.g., pride) or attitudes (e.g., religiosity); *locationist* perspectives on such functions, feelings, and attitudes and their hypothesized unique "signature" in the brain increasingly turn out to be wrong (e.g., *Bartholomew2004,Poldrack2006*). Likewise, as Lindquist et al. convincingly argue, emotions are not recognized by the brain as separate entities and, as such, do not each have their own seat and unique activation signature in the brain. Instead, Lindquist et al. present a strong case for a *constructionist* perspective in which emotions are comprised of multiple, more basic processes, which are each associated with their own location and activation signature in the brain. The combined outcomes of these processes result in the individual experience of a particular emotion.

If Lindquist et al.'s constructionist perspective is an accurate representation of the relation between emotions and the brain, what ramifications might this have for those psychiatric disorders that have emotional disturbances as core elements? Among other processes, Lindquist et al. distinguish between *core affect* (i.e., mental representation of bodily representations) and *conceptualization* (i.e., sensations from the body or external

Adapted from: Cramer, A. O. J., Kendler, K. S. & Borsboom, D. (2012). A constructionist account of emotional disorders. *Behavioral and Brain Sciences, 35*, 146-147.

world that are made meaningful). Major depression (MD) is a psychiatric disorder with "sadness" as one of the core elements, and it is well known that, in many cases, an episode of MD is preceded by stressful life events such as marital of health problems (e.g., Kendler et al., 1999). Although such life events are potentially quite aversive in nature, most people do not develop an episode of MD after experiencing them: So why are some people so severe affected by a stressful life event whereas most others are not? One explanation could be that in people who develop an episode of MD after a stressful life event the *conceptualization* process is dysfunctional; most people would respond with some sadness after a quarrel with a spouse (i.e., "normal" core affect), but in people with MD, this event is overly negatively conceptualized ("See, even my husband does not love me"). Such a hypothesis is consistent with clinical observations that patients with MD often engage in excessive rumination about past events (Roelofs, Huibers, Peeters, & Arntz, 2008; Roelofs, Huibers, Peeters, Arntz, & van Os, 2008).

On the other hand, in disorders with "fear" as a core element, the *core affect* process might have gone awry. Patients with a specific phobia are extremely fearful of certain objects (e.g., hypodermic needles), situations (e.g., flying an airplane), or animals (e.g., spiders) that do not elicit the same response in most other people. When confronted with, for example, a spider, patients with a phobia for that animal will respond with various bodily sensations (e.g., profuse sweating, heart palpitations) to that animal, whereas people without the phobia will not experience such bodily sensations; in terms of the Lindquist et al. perspective, specific phobia patents react with excessive *core affect* to phobic objects compared to non-phobic patients.

Distinguishing emotional disorders in terms of Lindquist et al.'s proposed processes might implicate a shift in clinical neuroscience from searching for *the* dysfunctional brain area causing a particular disorder to searching which brain areas do not *optimally work together* in perceiving and interpreting external stimuli (e.g., will we find that the conceptualization network is overly active in patents with MD?). This implication of Lindquist et al.'s work, that psychiatric disorders are not likely to be explained in terms of one dysfunctional brain area, bears a striking resemblance to recent attempts at formulating psychiatric disorders as networks of causally related symptoms (Borsboom, 2008b; Cramer et al., 2010; Kendler et al., 2011). In the *network approach*, psychiatric disorders are hypothesized to stem from direct interactions between symptoms (e.g., feeling tired $\rightarrow$ sleeping a lot $\rightarrow$ concentration problems) instead of from one underlying biological dysfunction (e.g., serotonin depletion causes all symptoms of MD). As such, each symptom is an autonomous causal entity and it is unlikely that such entities share the exact same etiological mechanisms. For example, symptoms such as insomnia and fatigue are likely governed by homeostatic processes, whereas symptoms such as guilty feelings and depressed mood are more likely regulated by cognitive processes (e.g., rumination). This hypothesis also lies at the heart of a theory in which psychiatric disorders are *mechanistic property clusters* (MPCs): mutually reinforcing networks of causal mechanisms at multiple levels of explanation (e.g., symptoms, brain). Each of these conceptualizations suggests that there are no hard delineations between disorders, as the processes that carry forward disturbances in a network are unlikely to be confined to a single set of symptoms (i.e., have a transdiagnostic character).

Thus, Lindquist et al.'s constructionist account is suggestive mutually reinforcing networks at the brain level that, when working optimally, result in the subjective experience of an appropriate particular emotion (e.g., fear when confronted with an angry grizzly bear). However, if one or more of those networks do not optimally work together, the result can be an inappropriate emotion (e.g., excessive fear when confronted with a spider). Subsequently, the network approach (i.e., mutually reinforcing networks at the symptom level) explains why, for example, a dysfunctional core affect process does not

result in a specific phobia (i.e., the entire syndrome) but results in excessive fear of a particular object or situation: other symptoms of a specific phobia, for example avoiding the feared object or situation, are a result of the excessive fear (i.e., one symptom causing the other). One way to investigate this hypothesis is by gathering intensive time series data with which one can accurately monitor the development of symptoms (and interactions among them) over time. This approach can be combined with frequent fMRI scans in order to link, for example, excessive activation of the conceptualization network, to the subsequent development of MD symptoms.

Appendix E

Validity from a network perspective

Abstract

This commentary was written in response to a target article by Newton (2012, published in the same issue of Measurement) that provided a comprehensive review of the history of validity as well as a proposal for a clearer and more precise definition of validity. In this commentary, I elaborate on how the definition of validity might change, more dramatically than proposed by Newton, when adopting a network stance towards certain psychological phenomena.

What is validity? A simple question but apparently one with many answers, as Newton highlights in his review of the history of validity. The current definition of validity, as entertained in the 1999 *Standards for Educational and Psychological Testing* is indeed a consensus, one between the classical notion of attributes, and measures thereof, and Cronbach and Meehl's (1955) proposition that validity should be concerned with interpretations of test scores and not with the tests themselves. Newton is certainly right when striving for a clearer and more precise definition of validity than the *Standards*' definition; and in many ways Newton succeeds in achieving this goal.

In both Newton's proposed clarification as well as the current *Standards*' definition, validity is closely associated with *measurement* of certain *attributes*. For many researchers in psychology, these terms are intimately associated with the measurement model that underlies the generic latent variable model: for example, insomnia (inso), fatigue (fati), concentration problems (conc) and depressed mood (mood) are caused by the latent attribute major depression (MD). Likewise, intelligence is hypothesized to be the *common cause* of people's responses to the items of an intelligence test. As such, when assessing the validity of a particular instrument that purportedly measures a single attribute, a confirmatory one-factor model is often part of the methodological toolkit: if the model fits, this is taken as evidence for the construct validity of that particular instrument (e.g., physiological hyperarousal: Joiner et al., 1999; psychopathy: Patrick, Edens, Poythress, Lilienfeld, & Benning, 2006). Importantly, for quite a few scholars in applied psychology research, these latent attributes—which, in a latent variable model, are nothing more than mathematical abstractions in a set of equations—are (implicitly) endowed with a

Adapted from: Cramer, A. O. J. (2012). Why the item "23 + 1" is not in a depression questionnaire: Validity from a network perspective. *Measurement, 10*, 50-54.

realist meaning: for example, personality psychologists who believe that the Extraversion factor in the Big Five factor model exists in the minds of individual people and causes behavior (McCrae & Costa, 2008).

In earlier work, we have postulated a new theoretical framework in which psychological phenomena like MD are *networks* in which the symptoms are not caused by a latent attribute but, instead, are directly related (Borsboom, 2008b; Cramer, Borsboom, et al., 2012; Cramer et al., 2010). That is, for example, the symptoms insomnia and fatigue are not caused by the latent "MD" attribute but, instead, are directly linked: insomnia $\rightarrow$ fatigue. Direct links need not be causal, they can be bidirectional as well (e.g., a bidirectional relation between depressed mood and concentration problems: depressed mood $\leftrightarrow$ concentration problems). Such causal/bidirectional relations between symptoms make more sense: Why would one need a latent attribute to explain why not sleeping and being tired are highly correlated? A similar argument was made for intelligence, in which, for example, verbal and arithmetic skills are strongly correlated because these skills mutually influence one another, and not because both skills are caused by the same underlying intelligence attribute (van der Maas et al., 2006). Finally, why would one need a latent "extraversion" attribute to explain why people who like parties often like to be in the centre of attention (Cramer, van der Sluis, et al., 2012).

In a classic measurement model, the probability of a symptom to be present depends on the latent attribute: the more depressed one is, the higher the probability that depression symptoms will be endorsed. In a network of symptoms, the probability of symptom x to be present depends on whether the symptoms that are directly linked to symptom x were present at an earlier point in time: for example, the probability of having concentration problems at $t = 1$ depends on whether sleep problems, fatigue, and worrying were present at $t = 0$. Or, in the case of extraversion, the probability of liking parties depends on whether one likes people and likes being in the center of attention.

Now, suppose the network perspective paints an accurate picture of at least some psychological phenomena. What does this mean for current definitions of validity as defined in terms of attributes and measurement? There is no attribute, at least not an attribute as defined in the classic measurement model. This does, however, not invalidate current definitions of validity. Take the network in Figure E.1 as an example and suppose we want to know the probability of being tired (fati). As mentioned earlier, this probability depends on whether the other symptoms with which fatigue is linked are present in an earlier point in time: the more depressed mood (mood), insomnia (inso), and concentration problems (conc), the higher the probability of fatigue to become present later in time. As such, it makes sense to say that fatigue in this network is *caused* by depressed mood, insomnia, and concentration problems; likewise, it makes sense to say that insomnia, in this network, is caused by depressed mood, fatigue, and concentration problems and so on. That is, each symptom in this network is caused by the *total activation of the other symptoms*; and it makes perfect sense to define the total activation summed over all symptoms to be an attribute —not an attribute in the classical measurement sense (with realist connotation), but one that is a summary statistic for how symptoms are influenced by one another.

So, from a network perspective, an intelligence test does not measure latent g but it assesses the extent to which cognitive processes such as verbal abilities and arithmetic skills are activated/present. Likewise, a depression questionnaire does not measure a latent depression attribute, but it assesses the total amount of activation of symptoms in a depression network. But to what does this total activation amount? In other work, we have formulated the hypothesis that psychological disorders are complex systems that, governed by a set of (non)linear differential equations, move towards one or more attractors (Schmittmann et al., 2013). For example, the MD network moves towards one

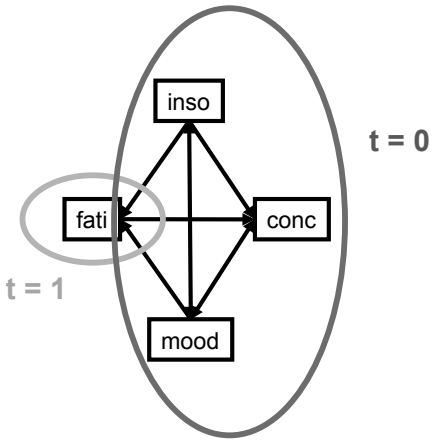

Figure E.1: A network of four symptoms. The probability of *fati* to be present at $t = 1$ depends on whether *inso*, *mood* and *conc* are present at an earlier point in time. *inso* = insomnia; *fati* = fatigue; *mood* = depressed mood; *conc* = concentration problems.

of two attractors: a "depressed" and a "non-depressed" attractor. Whenever the MD network is, for example, in the "depressed" attractor, we say that the MD network is in a depressed state. And what we do when administering a depression questionnaire to someone is assess the *state* of the depression network: the more symptoms are present/active, the more the MD network is pushed towards a depressed state.

Does this novel formulation of attributes and measurements mean that validity can no longer be assessed with, for example, fitting a confirmatory factor model? No. One can fit a one-factor model to investigate a questionnaire that consists of, say, extraversion items. However, when the model fits, the interpretation should be different: no longer should one speak of "construct validity" in the sense that the fit is evidence of an extraversion construct (with a possible realist connotation) and measurable by the items in that particular questionnaire. Rather, from a network perspective, the interpretation should be that the fit provides evidence that these particular items have direct nontrivial links with one another: apparently, endorsing an item depends on endorsing other items. As such, the latent factor serves as the summary statistic of the summed total activation/presence of these items.

So what is validity then, when adhering to a network perspective? Departing from Newton's as well as the *Standards*' definition, validity should be about the degree to which an instrument measures what it purports to measure (Ruch, 1924; see also Borsboom, Cramer, et al., 2009). And, then, what do instruments generally purport to measure? From a network perspective, instruments are designed to assess whether a particular set of items form a network, that is, whether they are directly connected with one another. Although it might seem to be so, this perspective on validity does not radically depart from existing formulations. In fact, in many cases the outcome of a validation study will be the same - take for instance the items "23 + 1" and "feeling blue". No matter the shape of the looking glass, "classic measurement", or "network", the conclusion will be that an instrument containing these two items is not a valid instrument for measuring MD.

However, the question of why this is so has different answers. From a classic measurement perspective, the instrument is not valid because there is no latent attribute that can explain the majority of covariation (if any, for the correlation between these items

will be low) between the two items. That is, "23 + 1" *does not measure* MD. From a network perspective, the instrument is not valid because the two items are not directly linked: the probability of correctly answering "23 + 1"does not depend on endorsing "feeling blue" and vice versa. One will conclude that the item "23 + 1" *is not part* of the MD *network*. That is, the interpretational difference between the two perspectives lies in assuming a measurement (i.e., classic measurement: the items measure an attribute), versus a mereological relationship, between attributes and items: from a network perspective, the items *are* what constitutes an attribute. Thus, in sum, feeling blue and thinking about suicide are not measurements of depression; feeling blue and thinking about suicide is what being depressed *is*.

References

Aan het Rot, M., Hogenelst, K., & Schoevers, R. A. (2012). Mood disorders in everyday life: a systematic review of experience sampling and ecological momentary assessment studies. *Clinical Psychology Reviews*, *32*, 510-523.

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, *96*, 358-372.

Achermann, P. (2004). The two-process model of sleep regulation revisited. *Aviation, Space, and Environmental Medicine*, *75*, A37-A43.

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for boltzmann machines. *Cognitive Science*, *9*, 147-169.

Aggen, S. H. (2005). Dsm criteria for major depression: Evaluating symptom patterns using latent-trait item response models. *Psychological Medicine*, *35*, 475-87.

Agresti, A. (2002). *Categorical data analysis*. Wiley.

Albert, R., & Barábasi, A. L. (1999). Emergence of scaling in random networks. *Science*, *286*, 509-512.

Albert, R., & Barábasi, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*, 47-97.

Albert, U., Rosso, G., Maina, G., & Bogetto, F. (2008). Impact of anxiety disorder comorbidity on quality of life in euthymic bipolar disorder patients: Differences between bipolair i and ii subtypes. *Journal of Affective Disorders*, *105*, 297-303.

Alligood, K. T., Sauer, T., & Yorke, J. A. (1997). *Chaos: An introduction to dynamical systems*. Springer.

Allport, G. W. (1946). Personalistic psychology as science: A reply. *Psychological Review*, *53*, 132-135.

Allport, G. W. (1960). The open system in personality theory. *Journal of Abnormal and Social Psychology*, *61*, 301-310.

Allport, G. W. (1968). The historical background of modern social psychology. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (Vol. 1). Addison-Wesley.

Amin, N., Schuur, M., Gusareva, E. S., Isaacs, A., Aulchenko, Y. S., Kirichenko, A. V., ... van Duijn, C. M. (2011). A genome-wide linkage study of individuals with high scores on neo personality traits. *Molecular Psychiatry*, *17*, 1031-1041.

Andersen, E. B. (1973). Sufficient statistics and latent trait models. *Psychometrika*, *42*, 69-81.

Anderson, J. C., Williams, S., McGee, R., & Silva, P. A. (1987). Dsm-iii disorders in preadolescent children: Prevalence in a large sample from the general population. *Archives of General Psychiatry*, *44*, 69-77.

Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry*, *40*, 57-87.

APA. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). American Psychiatric Association/APA Books.

Baldwin, G., & Shean, G. D. (2006). A taxometric study of the center for epidemiological studies depression scale. *Genetic, Social and General Psychology Monographs*, *132*, 101-128.

Barábasi, A. L. (2011). The network takeover. *Nature Physics*, *8*, 14-16.

Barlow, D. H., Allen, L. B., & Choate, M. L. (2004). Toward a unified treatment for emotional disorders. *Behavior Therapy*, *35*, 205-230.

Barrat, A., Barthélemy, M., & Vespignani, A. (2008). *Dynamical processe on complex networks*. Cambridge University Press.

Barrat, A., & Weigt, M. (2000). On the properties of small-world network models. *European Physical Journal B*, *13*, 547-460.

Bauer, M., Grof, P., Rasgon, N., Bschor, T., Glenn, T., & Whybrow, P. C. (2006). Temporal relation between sleep and mood in patients with bipolar disorder. *Bipolar Disorders*, *8*, 160-167.

Bebbington, P. (1987). Misery and beyond: the pursuit of disease theories of depression. *International Journal of Social Psychiatry*, *33*, 13-20.

Beck, A. T. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy*, *1*, 5-38.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. Guilford Press.

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: differential susceptibility to environmental influences. *Psychological Bulletin*, *135*, 885-908.

Berg, M. d., Cheong, O., Kreveld, M., & Overmars, M. (2008). *Computational geometry: algorithms and applications*. Springer-Verlag.

Berrios, G. E. (1993). European views on personality disorders: a conceptual history. *Comprehensive Psychiatry*, *34*, 14-30.

Biederman, J., Faraone, S., Mick, E., & Lelon, E. (1995). Psychiatric comorbidity among referred juveniles with major depression: Fact or artifact? *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*, 579-590.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (p. 397-479). Addison-Wesley.

Birnbaum, K. (1923). *Der aufbau der psychose*. Springer.

Bleich, A., Koslowsky, M., Dolev, A., & Lerer, R. (1997). Post-traumatic stress disorder and depression: An analysis of comorbidity. *British Journal of Psychiatry*, *170*, 479-482.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*, 175-308.

Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*, 605-634.

Boomsma, D. I., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics*, *3*, 872-882.

Borbély, A. A., & Achermann, P. (1999). Sleep homeostasis and models of sleep regulation. *Journal of Biological Rhythms*, *14*, 559-568.

Bornstein, R. F. (2003). Behaviorally referenced experimentation and symptom validation: A paradigm for 21st-century personality disorder research. *Journal of Personality Disorders*, *17*, 1-18.

Borsboom, D. (2002). Letter to the editor. the structure of the dsm. *Archives of General Psychiatry*, *59*, 569-570.

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge University Press.

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, *71*, 425-440.

Borsboom, D. (2008a). Latent variable theory. *Measurement: Interdisciplinary Research and Perspective*, *6*, 25-53.

Borsboom, D. (2008b). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, *64*, 1089-1108.

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*.

Borsboom, D., Cramer, A. O. J., Kievit, R. A., Zand Scholten, A., & Franic, S. (2009). The end of construct validity. In R. W. Lissitz (Ed.), *The concept of validity.* Information Age Publishers.

Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The small world of psychopathology. *PLoS ONE*, *6*, e27407.

Borsboom, D., Kievit, R. A., Cervone, D. P., & Hood, S. B. (2009). The two disciplines of scientific psychology, or: the disunity of psychology as a working hypothesis. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudary (Eds.), *Developmental process methodology in the social and developmental sciences.* Springer.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203-219.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*, 1061-1071.

Bouchard, T. J. (1994). Genes, environment, and personality. *Science*, *264*, 1700-1701.

Boyd, J. H., Burke, J. D., Gruenberg, E., Holzer, C. E., Rae, D. S., George, L. K., ... Nestadt, G. (1984). Exclusion criteria of dsm-iii: A study of co-occurrence of hierarchy-free syndromes. *Archives of General Psychiatry*, *41*, 983-989.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, *25*, 163-177.

Bremner, J. D., Narayan, M., Anderson, E. R., Staib, L. H., Miller, H. L., & Charney, D. S. (2000). Hippocampal volume reduction in major depression. *American Journal of Psychiatry*, *157*, 115-117.

Brown, G. W., Bifulco, A., & Harris, T. O. (1987). Life events, vulnerability and onset of depression: some refinements. *British Journal of Psychiatry*, *150*, 30-42.

Brown, T. A., Antony, M. M., & Barlow, D. H. (1995). Diagnostic comorbidity in panic disorder: Effect on treatment outcome and course of comorbid diagnoses following treatment. *Journal of Consulting and Clinical Psychology*, *63*, 408-418.

Brown, T. A., & Barlow, D. H. (1992). Comorbidity among anxiety disorders: Implications for treatment and dsm-iv. *Journal of Consulting and Clinical Psychology*, *60*, 835-844.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the dsm-iv anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology*, *110*, 585-599.

Brown, T. A., Chorpita, B. F., & Barlow, D. H. (1998). Structured relationships among dimensions of the dsm-iv anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal. *Journal of Abnormal Psychology*, *107*, 179-192.

Bunney, B. G., & Bunney, W. E. (2012). Mechanisms of rapid antidepressant effects of sleep deprivation therapy: Clock genes and circadian rhythms. *Biological Psychiatry*.

Caprara, G. V., & Cervone, D. (2000). *Personality: determinants, dynamics, and potentials*. Cambridge University Press.

Carpenter, S. R., & Brock, W. A. (2006). Rising variance: a leading indicator of ecological transition. *Ecology Letters*, *9*, 311-318.

Carter, R. M., Wittchen, H. U., Pfister, H., & Kessler, R. C. (2001). One-year prevalence of subthreshold and threshold dsm-iv generalized anxiety disorder in a nationally representative sample. *Depression and Anxiety*, *13*, 78-88.

Caspi, A., Bem, D. J., & Elder, G. H. (1989). Continuities and consequences of interactional styles across the life course. *Journal of Personality*, *57*, 375-406.

Caspi, A., Elder, G. H., & Bem, D. J. (1987). Moving against the world: life-course patterns of explosive children. *Developmental Psychology*, *23*, 308-313.

Caspi, A., Elder, G. H., & Bem, D. J. (1988). Moving away from the world: life-course patterns of shy children. *Developmental Psychology*, *24*, 824-831.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., ... Poulton, R. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-htt gene. *Science*, *301*, 386-389.

Cattell, R. B. (1950). *Personality: A systematic theoretical and factual study*. McGraw-Hill.

Cervone, D. (2004). The architecture of personality. *Psychological Review*, *111*(183-204).

Cervone, D. (2005). Personality architecture: within-person structures and processes. *Annual Review of Psychology*, *56*, 423-452.

Cilliers, P. (1998). *Complexity and postmodernism*. Routledge.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P., & Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation and

imipramine in the treatment of panic disorder. *British Journal of Psychiatry*, *164*, 759-769.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, *100*, 316-336.

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology*, *46*, 121-153.

Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract or not abstract? *Behavioral and Brain Sciences*, *32*, 313-373.

Coryell, W., Winokur, G., Shea, T., Maser, J. D., Endicott, J., & Akiskal, H. S. (1994). The long-term stability of depressive subtypes. *American Journal of Psychiatry*, *151*, 199-204.

Cox, D. R., & Wermuth, N. (1993). Linear dependencies represented by chain graphs. *Statistical Science*, *8*, 204-218.

Cramer, A. O. J. (2012). Why the item "23 + 1" is not in a depression questionnaire: Validity from a network perspective. *Measurement: Interdisciplinary Research and Perspective*, *10*, 50-54.

Cramer, A. O. J., Borsboom, D., Aggen, S. H., & Kendler, K. S. (2012). The pathoplasticity of dysphoric episodes: differential impact of stressful life events on the pattern of depressive symptom inter-correlations. *Psychological Medicine*, *42*, 957-965.

Cramer, A. O. J., Kendler, K. S., & Borsboom, D. (2011). Where are the genes? the implications of a network perspective on gene hunting in psychopathology. *European Journal of Personality*, *25*, 270-271.

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., ... Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, *26*, 414-431.

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*, 137-193.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.

Csárdi, G., & Nepusz, T. (2006). The igraph software package for complex networks research. *International Journal of Complex Systems*, 1695.

Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., & Held, H. (2008). Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences USA*, *105*, 14308-14312.

Danks, D., Fancsali, S., Glymour, C., & Scheines, R. (2010). Comorbid science? *Behavioral and Brain Sciences*, *33*, 153-154.

Darwin, C. (1871). *The descent of man, and selection in relation to sex.* John Murray.

David, M., Ceschi, G., Billieux, J., & van der Linden, M. (2008). Depressive symptoms after trauma. is self-esteem a mediating factor? *The Journal of Nervous and Mental Disease*, *196*, 735-742.

de Moor, M. H. M., Costa, P. T., Terracciano, A., Krueger, R. F., de Geus, E. J. C., Toshiko, T., ... Boomsma, D. I. (2010). Meta-analysis of genome-wide association studies for personality. *Molecular Psychiatry*, *17*, 337-349.

DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., ... Gallop, R. (2005). Cognitive therapy vs medications in the treatment of moderate to severe depression. *Archives of General Psychiatry*, *62*, 409-416.

Dethlefsen, D., & Hojsgaard, S. (2005). A common platform for graphical models in r: The grbase package. *Journal of Statistical Software*, *15*, 1-12.

DiBattista, G., Eades, P., Tamassia, R., & Tollis, I. G. (1994). Algorithms for drawing graphs: an annotated bibliography. *Computational Geometry: Theory and Applications*, *4*, 235-282.

Digman, J. M. (1990). Personality structure: emergence of the five-factor model. *Annual Review of Psychology*, *41*, 417-440.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerical Mathematics*, *1*, 269-271.

Diks, C., & Panchenko, V. (2005). A note on the hiemstra-jones test for granger non-causality. *Studies in Nonlinear Dynamics and Econometrics*, *9*.

Diks, C., & Panchenko, V. (2006). A new statistic and practical guidelines for nonparametric granger causality testing. *Journal of Economic Dynamics and Control*, *30*, 1647-1669.

Dolan, C. V., Oort, F. J., Stoel, R. D., & Wicherts, J. M. (2009). Testing measurement invariance in the target rotated multigroup exploratory factor model. *Structural Equation Modeling*, *16*, 295-314.

Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology*, *25*, 117-129.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*, 155-174.

Ehlers, C. L. (1995). Chaos and complexity. can it help us understand mood and behavior? *Archives of General Psychiatry*, 960-964.

Eichler, M. (2007). Granger causality and path diagrams for multivariate time series. *Journal of Econometrics*, *137*, 334-353.

Ellis, J. L., & Junker, B. W. (1997). Tail-measurability in monotone latent variable models. *Psychometrika*, *62*, 495-523.

Ensel, W. M., & Lin, N. (1996). Distal stressors and the life stress process. *Journal of Community Psychology*, *24*, 66-82.

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*, 1-18.

Erdös, P., & Rényi, A. (1959). On random graphs. *Publications of Mathematics Debrecen*, *6*, 290-297.

Eysenck, H. J. (1967). *The biological basis of personality.* Thomas.

Finelli, L. A., Baumann, H., Borbély, A. A., & Achermann, P. (2000). Dual electroencephalogram markers of human sleep homeostasis: correlation between theta activity in waking and slow-wave activity in sleep. *Neuroscience*, *101*, 523-529.

Flay, B. R. (1978). Catastrophe theory in social psychology: some applications to attitudes and social behavior. *Behavior Sciences*, *23*, 335-350.

Flett, G. L., Vredenburg, K., & Krames, L. (1997). The continuity of depression in clinical and nonclinical samples. *Psychological Bulletin*, *121*, 395-416.

Fort, H., Mazzeo, N., Scheffer, M., & van Nes, E. H. (2010). Catastrophic shifts in ecosystems: spatial early warnings and management procedures. *Journal of Physics Conference Series*, *246*, 1-15.

Franklin, C. L., & Zimmerman, M. (2001). Posttraumatic stress disorder and major depressive disorder: Investigating the role of overlapping symptoms in diagnostic comorbidity. *The Journal of Nervous and Mental Disease*, *189*, 548-551.

Franklin, M. E., & Foa, E. B. (2011). Treatment of obsessive compulsive disorder. *Annual Review of Clinical Psychology*, *7*, 229-243.

Frazer, A., & Benmansour, S. (2002). Delayed pharmacological effects of antidepressants. *Molecular Psychiatry*, *7*, S23-S28.

Frazer, K. A., Murray, S. S., Schork, N. J., & Topo, E. J. (2009). Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics*, *10*, 241-251.

Freitas, R. A. (2007). Personal choice in the coming era of nanomedicine. In F. Allhoff (Ed.), *Nanoethics: The ethical implications of nanotechnology.* John Wiley and Sons.

Frewen, P. A., Allen, S. L., Lanius, R. A., & Neufeld, R. W. J. (2012). Perceived causal relations: novel methodology for assessing client attributions about causal associations between variables including symptoms and functional impairment. *Assessment*, *19*, 480-493.

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software Practice and Experience*, *21*, 1129-1164.

Fullerton, J., Cubin, M., Tiwari, H., Wang, C., Bohmra, A., Davidson, S., ... Flint, J. (2003). Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative-trait loci that influence variation in the human personality trait neuroticism. *The American Journal of Human Genetics*, *72*, 879-890.

Funder, D. C. (1991). Global traits: A neo-allportian approach to personality. *Psychological Science*, *2*, 31-39.

Gall, F. J., & Spurzheim, J. C. (1835). *On the functions of the brain and each of its parts: With observations on the possibility of determining the instincts, propensities, and talents, or the moral and intellectual dispositions of man and animals, by the configuration of the brain and head.* Marsh, Capen, and Lyon.

Gamallo, P., Gasperin, C., Agustini, A., & Lopes, G. P. (2001). Syntactic-based methods for measuring word similarity. In *Text, speech and dialogue* (p. 116-125). Springer Berlin Heidelberg.

Geschwind, N., Peeters, F., Drukker, M., van Os, J., & Wichers, M. (2011). Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression: a randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *79*, 618-628.

Gillin, J. (1983). The sleep therapies of depression. *Progress in Neuropsychophar-*

macology and Biological Psychiatry, *7*, 351-364.

Gillin, J. C., Buchsbaum, M., Wu, J., Clark, C., & Bunney, W. (2001). Sleep deprivation as a model experimental antidepressant treatment: findings from functional brain imaging. *Depression and Anxiety*, *14*, 37-49.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*, 26-34.

Goldbeter, A. (2011). A model for the dynamics of bipolar disorders. *Progress in Biophysics and Molecular Biology*, *105*, 119-127.

Gossop, M., Stewart, D., Browne, N., & Marsden, J. (2002). Factors associated with abstinence, lapse or relapse to heroin use after residential treatment: Protective effect of coping responses. *Addiction*, *97*, 1259-1267.

Gotlib, I. H., Lewinsohn, P. M., & Seeley, J. R. (1995). Symptoms versus a diagnosis of depression: Differences in psychosocial functioning. *Journal of Consulting and Clinical Psychology*, *63*, 90-100.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*, 424-438.

Grayson, D. A. (1988). Two-group classicitaion in latent trait theory: scores with monotone likelihood ratio. *Psychometrika*, *53*, 383-392.

Greden, J. F. (2001). The burden of recurrent depression: Causes, consequences, and future prospects. *Journal of Clinical Psychiatry*, *62*, 5-9.

Greenberg, A. S., & Bailey, J. M. (1993). Do biological explanations of homosexuality have moral, legal, or policy implications? *The Journal of Sex Research*, *30*, 245-251.

Grimmett, G. (2010). *Probability on graphs. random processes on graphs and lattices*. Cambridge University Press.

Haig, B. (2005). An abductive theory of scientific method. *Psychological Methods*, *10*, 371-388.

Haig, B. (2009). Inference to the best explanation: A neglected approach to theory appraisal in psychology. *American Journal of Psychology*, *122*, 219-234.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: rationale and application of multivariate stationary time series. *Multivariate Behavior Research*, *40*, 207-233.

Hamaker, E. L., Nesselroade, J. R., & Molenaar, P. C. M. (2007). The integrated state-trait model. *Journal of Research in Personality*, *41*, 295-315.

Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology*, *1*, 293-319.

Hammen, C. L., & Peters, S. D. (1978). Interpersonal consequences of depression: Responses to men and women enacting a depressed role. *Journal of Abnormal Psychology*, *87*, 322-332.

Hare, R. D. (2003). *Manual for the revised psychopathy checklist* (2nd ed.). Multi-Health Systems.

Harman, H. H. (1960). *Modern factor analysis*. University of Chicago Press.

Harmer, C. J. (2010). Antidepressant drug action: a neuropsychological perspective. *Depression and Anxiety*, *27*, 231-233.

Harmer, C. J., Shelley, N. C., Cowen, P. J., & Goodwin, G. M. (2004). Increased positive versus negative affective perception and memory in healthy volun-

teers following selective serotonin and norepinephrine reuptake inhibition. *American Journal of Psychiatry*, *161*, 1256-1263.

Harris, T. O., Borsanyi, S., Messari, S., Stanford, K., Brown, G. W., Cleary, S. E., ... Herbert, J. (2000). Morning cortisol as a risk factor for subsequent major depressive disorder in adult women. *British Journal of Psychiatry*, *177*, 505-510.

Hartman, C., & Benes, B. (2006). Autonomous boids. *Computer Animation and Virtual Worlds*, *17*, 199-206.

Hartman, C. A., Hox, J., Mellenbergh, G. J., Boyle, M. H., Offord, D. R., Racine, Y., ... Sergeant, J. A. (2001). Dsm-iv internal construct validity: When a taxonomy meets data. *Journal of Child Psychology and Psychiatry*, *42*, 817-836.

Haslam, N. (2000). Psychiatric categories as natural kinds: Essentialist thinking about mental disorders. *Social Research*, *67*, 1031-1058.

Haslam, N., & Ernst, D. (2002). Essentialist beliefs about mental disorders. *Journal of Social and Clinical Psychology*, *21*, 628-644.

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychological Medicine*, *42*, 903-920.

Heady, B., & Wearing, A. (1989). Personality, life events, and subjective well-being: toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*, *57*, 731-739.

Hemmeter, U., Hemmeter-Spernal, J., & Krieg, J. (2010). Sleep deprivation in depression. *Expert Review of Neurotherapeutics*, *10*, 1101-1115.

Henry, B., Moffitt, T. E., Caspi, A., Langley, J., & Silva, P. A. (1994). On the "remembrance of things past": a longitudinal evaluation of the retrospective method. *Psychological Assessment*, *6*, 92-101.

Herman, I. (2000). Graph visualization and navigation in information visualization: a survey. *IEEE Transactions on Visualization and Computer Graphics*, *6*, 24-43.

Hettema, J. M., Neale, M. C., Myers, J. M., Prescott, C. A., & Kendler, K. S. (2006). A population-based twin study of the relationship between neuroticism and internalizing disorders. *American Journal of Psychiatry*, *163*, 857-864.

Hiemstra, C., & Jones, J. D. (1994). Testing for linear and nonlinear granger causality in the stock price-volume relation. *Journal of Finance*, *49*, 1639-1664.

Hirshfeld, D. R., Rosenbaum, J. F., Biederman, J., Bolduc, E. A., Faraone, S. V., Snidman, N., ... Kagan, J. (1992). Stable behavioral inhibition and its association with anxiety disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 103-111.

Hoffman, C. D., Sweeney, D. P., Lopez-Wagner, M. C., Hodge, D., Nam, C. Y., & Botts, B. H. (2008). Children with autism: sleep problems and mothers' stress. *Focus Autism and Other Developmental Disabilities*, *23*, 155-165.

Holland, P. W., & Rosenbaum, P. R. (1986). Conditional association and unidimensionality in monotone latent variable models. *Annals of Statistics*, *14*, 1523-1543.

Hood, S. B. (2008). *Latent variable realism in psychometrics*. Unpublished doctoral dissertation, Indiana University.

Howes, O. D., & Kapur, S. (2009). The dopamine hypothesis of schizophrenia: Version iii - the final common pathway. *Schizophrenia Bulletin*, *35*, 549-562.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Hubert, L. J. (1994). Some applications of graph theory to clustering. *Psychometrika*, *39*, 283-309.

Hubert, M. T., Braun, H. A., & Krieg, J. C. (1999). Consequences of deterministic and random dynamics for the course of affective disorders. *Biological Psychiatry*, *46*, 256-262.

Hyland, M. E. (2011). *The origins of health and disease*. Cambridge University Press.

Ising, E. (1925). Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, *31*, 253-258.

Jablensky, A. (2007). Does psychiatry need an overarching concept of "mental disorder"? *World Psychiatry*, *6*, 157-158.

Jacobs, N., Kenis, G., Peeters, F., Derom, C., Vlietinck, R., & van Os, J. (2006). Stress-related negative affectivity and genetically altered serotonin transporter function. *Archives of General Psychiatry*, *63*, 989-996.

Jang, K. L., Livesley, W. J., & Vernon, P. A. (1996). Heritability of the big five personality dimensions and their facets: a twin study. *Journal of Personality*, *64*, 577-591.

Joiner, T. E., Steer, R. A., Beck, A. T., Schmidt, N. B., Rudd, M. D., & Catanzaro, S. J. (1999). Physiological hyperarousal: Construct validity of a central aspect of the tripartite model of depression and anxiety. *Journal of Abnormal Psychology*, *108*, 290-298.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, *36*, 109-133.

Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, *25*, 258-272.

Kalisch, M., Maechler, M., Colombo, D., Maathuis, M. H., & Buehlmann, P. (2012). Causal inference using graphical models with the r package pcalg. *Journal of Statistical Software*, *47*, 1-26.

Kamphuis, J. H., & Telch, M. J. (2000). Effects of distraction and guided threat reappraisal on fear reduction during exposure-based treatment for specific fears. *Behaviour Research and Therapy*, *38*, 1163-1181.

Kandel, E. R., & Squire, L. R. (1992). Cognitive neurosciences: Editorial overview. *Current Opinion in Neurobiology*, *2*, 143-145.

Kashani, J. H., Beck, N. C., Hoeper, E. W., Fallahi, C., Corcoran, C. M., McAllister, J. A., ... Reid, J. C. (1987). Psychiatric disorders in a community sample of adolescents. *American Journal of Psychiatry*, *144*, 584-589.

Keller, M. C., Neale, M. C., & Kendler, K. S. (2007). Association of different adverse life events with distinct patterns of depressive symptoms. *American Journal of Psychiatry*, *164*, 1521-1529.

Keller, M. C., & Nesse, R. M. (2005). Subtypes of low mood provide evidence of its adaptive significance. *Journal of Affective Disorders*, *86*, 27-35.

Keller, M. C., & Nesse, R. M. (2006). The evolutionary significance of depressive symptoms: different adverse situations lead to different depressive symtom patterns. *Journal of Personality and Social Psychology*, *91*, 316-330.

Kendell, R., & Jablensky, A. (2003). Distinguishing between the validity and utility of psychiatric diagnoses. *American Journal of Psychiatry*, *160*, 4-12.

Kendell, R. E. (1975). The concept of disease and its implications for psychiatry. *British Journal of Psychiatry*, *127*, 305-315.

Kendler, K. S. (2005a). "a gene for...": the nature of gene action in psychiatric disorders. *American Journal of Psychiatry*, *162*, 1243-1252.

Kendler, K. S. (2005b). Toward a philosophical structure for psychiatry. *American Journal of Psychiatry*, *162*, 433-440.

Kendler, K. S. (2006). Reflections on the relationship between psychiatric genetics and psychiatric nosology. *American Journal of Psychiatry*, *163*, 1138-1146.

Kendler, K. S., & Baker, J. H. (2007). Genetic influences on measures of the environment: a systematic review. *Psychological Medicine*, *37*, 615-626.

Kendler, K. S., Gardner, C. O., Neale, M. C., & Prescott, C. A. (2001). Genetic risk factors for major depression in men and women: similar or distinct heritabilities and same or partly distinct genes. *Psychological Medicine*, *48*, 191-214.

Kendler, K. S., Gardner, C. O., & Prescott, C. A. (2003). Personality and the experience of environmental adversity. *Psychological Medicine*, *33*, 1193-1202.

Kendler, K. S., Gatz, M., Gardner, C. O., & Pedersen, N. L. (2006). A swedish national twin study of lifetime major depression. *American Journal of Psychiatry*, *163*, 109-114.

Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1999). Causal relationships between stressful life events and the onset of major depression. *American Journal of Psychiatry*, *156*, 837-841.

Kendler, K. S., Kessler, R. C., Walters, E. E., MacLean, C., Neale, M. C., Heath, A. C., & Eaves, L. J. (1995). Stressful life events, genetic liability, and onset of an episode of major depression in women. *American Journal of Psychiatry*, *152*, 833-842.

Kendler, K. S., & Myers, J. (2010). The genetic and environmental relationship between major depression and the five-factor model of personality. *Psychological Medicine*, *40*, 801-806.

Kendler, K. S., & Prescott, C. A. (2006). *Genes, environment, and psychopathology: Understanding the causes of psychiatric and substance use disorders*. Guilford Press.

Kendler, K. S., Walters, E. E., & Kessler, R. C. (1997). The prediction of length of major depressive episodes: results from an epidemiological sample of female twins. *Psychological Medicine*, *27*, 107-117.

Kendler, K. S., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine*, *41*, 1143-1150.

Kennedy, N., & Paykel, E. S. (2004). Residual symptoms at remission from depression: Impact on long-term outcome. *Journal of Affective Disorders*,

80, 135-144.

Kenny, D. A., & Zautra, A. (1995). The trait-state-error model for multiwave data. *Journal of Consulting and Clinical Psychology*, *63*, 52-59.

Kessler, R. C. (1997). The effects of stressful life events on depression. *Annual Review of Psychology*, *48*, 191-214.

Kessler, R. C., Berglund, P., Chiu, W. T., Demler, O., Heeringa, S., Hiripi, E., ... Zheng, H. (2004). The us national comorbidity survey replication (ncs-r): Design and field procedures. *International Journal of Methods in Psychiatric Research*, *13*, 69-92.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of dsm-iv disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, *62*, 593-602.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month dsm-iv disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, *62*, 617-627.

Kessler, R. C., DuPont, R. L., Berglund, P., & Wittchen, H. U. (1999). Impairment in pure and comorbid generalized anxiety disorder and major depression at 12 months in two national surveys. *American Journal of Psychiatry*, *156*, 1915-1923.

Kessler, R. C., & Magee, W. J. (1993). Childhood adversities and adult depression: basic patterns of associations in a us national survey. *Psychological Medicine*, *23*, 679-690.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., ... Kendler, K. S. (1994). Lifetime and 12-month prevalence of dsm-iii-r psychiatric disorders in the united states: Results from the national comorbidity survey. *Archives of General Psychiatry*, *51*, 8-19.

Kessler, R. C., & Ustun, T. B. (2004). The world mental health (wmh) survey initiative version of the world health organization (who) composite international diagnostic interview (cidi). *International Journal of Methods in Psychiatric Research*, *13*, 93-121.

Kievit, R. A., Romeijn, J. W., Waldorp, L. J., Wicherts, J. M., Scholte, S., & Borsboom, D. (2011). Modeling mind and matter: Reductionism and psychological measurement in cognitive neuroscience. *Psychological Inquiry*, *22*, 139-157.

Kim, N. S., & Ahn, W.-k. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology: General*, *131*, 451-476.

Klein, D. F. (1978). A proposed definition of mental disorder. In D. F. Klein & R. L. Spitzer (Eds.), *Critical issues in psychiatric diagnosis.* Raven Press.

Klein, D. F., & Davis, J. M. (1969). *Diagnosis and drug treatment of psychiatric disorders.* Williams and Wilkins.

Klein, D. N., Kotov, R., & Bufferd, S. J. (2011). Personality and depression: explanatory models and review of the evidence. *Annual Review of Clinical Psychology*, *7*, 269-295.

Kolaczyk, E. D. (2009). *Statistical analysis of network data.* Springer.

Kraepelin, E. (1921). *Manic-depressive insanity and paranoia.* E. and S. Livingstone.

Kraepelin, E. (1923). *Clinical psychiatry: A textbook for students and physicians.* MacMillan.

Krichel, T., & Bakkalbasi, N. (2006). A social network analysis of research collaboration in the economics community. *Journal of Information Management and Scientometrics*, *3*, 1-12.

Krøgsboll, L. T., Hróbjartsson, A., & Gøtzsche, P. C. (2009). Spontaneous improvement in randomised clinical trials: meta-analysis of three-armed trials comparing no treatment, placebo and active intervention. *BMC Medical Research Methodology*, *9*, 1.

Krueger, R. F. (1999). The structure of common mental disorders. *Archives of General Psychiatry*, *56*, 921-926.

Krueger, R. F., Caspi, A., Moffitt, T. E., Silva, P. A., & McGee, R. (1996). Personality traits are differentially linked to mental disorders: a multitrait-multidiagnosis study of an adolescent birth cohort. *Journal of Abnormal Psychology*, *105*, 299-312.

Krueger, R. F., & Markon, K. E. (2006). Understanding psychopathology. melding behavior genetics, personality, and quantitative psychology to develop an empirically based model. *Current Directions in Psychological Science*, *15*, 113-117.

Kuhs, H., & Tölle, R. (1991). Sleep deprivation therapy. *Biological Psychiatry*, *29*, 1129-1148.

Kuo, P. H., Neale, M. C., Riley, B. P., Patterson, D. G., Walsh, D., Prescott, C. A., & Kendler, K. S. (2007). A genome-wide linkage analysis for the personality trait neuroticism in the irish affected sib-pair study of alcohol dependence. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *144B*, 463-468.

Kuyken, W., Byford, S., Taylor, R. S., Watkins, E., Holden, E., White, K., ... Teasdale, J. D. (2008). Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. *Journal of Consulting and Clinical Psychology*, *76*, 966-978.

Lacasse, J. R., & Leo, J. (2005). Serotonin and depression: A disconnect between the advertisements and the scientific literature. *PLoS Medicine*, *2*, e392.

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8-9-year old students. *Cognition*, *93*, 99-125.

Landsness, E. C., Goldstein, M. R., Peterson, M. J., Tononi, G., & Benca, R. M. (2011). Antidepressant effects of selective slow wave sleep deprivation in major depression: A high-density eeg investigation. *Journal of Psychiatric Research*, *45*, 1019-1026.

Leibenluft, E., Albert, P. S., Rosenthal, N. E., & Wehr, T. A. (1996). Relationship between sleep and mood in patients with rapid-cycling bipolar disorder. *Psychiatry Research*, *63*, 161-168.

Leskelä, U. S., Melartin, T. K., Lestelä-Mielonen, P. S., Rystälä, H. J., Sokero, T. P., Heikkinen, M. E., & Isometsä, E. T. (2004). Life events, social support, and onset of major depressive episode in finnish patients. *Journal of Nervous*

and Mental Disease, *192*, 373-381.

Levandowsky, M., & Winter, D. (1971). Distance between sets. *Nature*, *234*, 34-35.

Levinson, D. F. (2006). The genetics of depression: A review. *Biological Psychiatry*, *60*, 84-92.

Lewis, A. J. (1934). Melancholia: A clinical survey of depressive states. *British Journal of Psychiatry*, *80*, 277-378.

Lilienfeld, S. O., & Marino, L. (1999). Essentialism revisited: Evolutionary theory and the concept of mental disorder. *Journal of Abnormal Psychology*, *108*, 400-411.

Loehlin, J. C. (1992). *Genes and environment in personality development*. Sage.

Lopez, A. D., Mathers, C. D., Ezzati, M., Jamison, D. T., & Murray, C. J. L. (2006). Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *The Lancet*, *367*, 1747-1757.

Lord, F. M. (1953). The relation of test score to the trait underlying the test. *Educational and Psychological Measurement*, *13*, 517-549.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley.

Low, N. C. P., Cui, L., & Merikangas, K. R. (2008). Specificity of familial transmission of anxiety and comorbid disorders. *Journal of Psychiatric Research*, *42*, 596-604.

Ma, S. H., & Teasdale, J. D. (2004). Mindfulness-based cognitive therapy for depression: Replication and exploration of differential relapse prevention effects. *Journal of Consulting and Clinical Psychology*, *72*, 31-40.

Maciejewski, P. K., Prigerson, H. G., & Mazure, C. M. (2000). Self-efficacy as a mediator between stressful life events and depressive symptoms: Differences based on history of prior depression. *British Journal of Psychiatry*, *176*, 373-378.

Mackey, M. C., & Glass, L. (1977). Oscillation and chaos in physiological control systems. *Science*, *197*, 287-289.

MacQueen, G., Campbell, S., McEwen, B. S., Macdonald, K., Amano, S., Joffe, R. T., ... Young, L. T. (2003). Course of illness, hippocampal function, and hippocampal volume in major depression. *Proceedings of the National Academy of Sciences USA*, *100*, 1387-1392.

MacQueen, G., & Frodl, T. (2011). The hippocampus in major depression: evidence for the convergence of the bench and bedside in psychiatric research? *Molecular Psychiatry*, *16*, 252-264.

Maes, H. H., Neale, M. C., Chen, X., Chen, J., Prescott, C. A., & Kendler, K. S. (2011). A twin association study of nicotine dependence with markers in the chrna3 and chrna5 genes. *Behavior Genetics*, *41*, 680-690.

Maher, B. (2008). Personal genomes: the case of the missing heritability. *Nature*, *456*, 18-21.

Maier, W., Gänsicke, M., & Weiffenbach, O. (1997). The relationship between major and subthreshold variants of unipolar depression. *Journal of Affective Disorders*, *45*, 41-51.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., ... Visscher, P. M. (2009). Finding the missing heritability of complex

diseases. *Nature*, *461*, 747-753.
Markowitsch, H. J., Vandekerckhove, M. M. P., Lanfermann, H., & Russ, M. O. (2003). Engagement of lateral and medial prefrontal areas in the ecphory of sad and happy autobiographical memories. *Cortex*, *39*, 643-665.
Markus, K. A. (2008). Hypothesis formulation, model interpretation, and model equivalence: Implications of a mereological causal interpretation of structural equation models. *Multivariate Behavior Research*, *43*, 177-209.
Maser, J. D., Kaelber, C., & Weise, R. E. (1991). International use and attitudes towards dsm-iii and dsm-iii-r: growing consensus in psychiatric classifications. *Journal of Abnormal Psychology*, *100*, 271-279.
Masters, G. N., & Wright, B. D. (1984). The essential process in a family of measurement models. *Psychometrika*, *49*, 529-544.
McCrae, R. R., & Costa, P. T. (1995). Trait explanations in personality psychology. *European Journal of Personality*, *9*, 231-252.
McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*, 509-516.
McCrae, R. R., & Costa, P. T. (2008). Empirical and theoretical status of the five-factor model of personality traits. In G. Boyle, G. Matthews, & D. Saklofske (Eds.), *Handbook of personality theory and assessment* (Vol. 1, p. 273-294). Sage.
McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised neo personality inventory: confirmatory factor analysis versus procrustes rotation. *Journal of Personality and Social Psychology*, *70*, 552-566.
McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology*, *34*, 100-117.
McGrath, R. E. (2005). Conceptual complexity and construct validity. *Journal of Personality Assessment*, *85*, 112-124.
McGue, M., & Christensen, K. (2003). The heritability of depression symptoms in elderly danish twins: occasion-specific versus general effects. *Behavior Genetics*, *33*, 83-93.
McGuffin, P., Katz, R., & Bebbington, P. (1988). The camberwell collaborative depression study iii. depression and adversity in the relatives of probands. *British Journal of Psychiatry*, *153*, 775-782.
McNally, R. J. (1994). *Panic disorder: A critical analysis*. Guilford Press.
McNiel, J. M., & Fleeson, W. (2006). The causal effects of extraversion on positive affect and neuroticism on negative affect: Manipulating state extraversion and state neuroticism in an experimental approach. *Journal of Research in Personality*, *40*, 529-550.
Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality*, *60*, 117-174.
Merikangas, K. R., Mehta, R. L., Molnar, B. E., Walters, E. E., Swendsen, J. D., Aguilar-Gaziola, S., ... Kessler, R. C. (1998). Comorbidity of substance use disorders with mood and anxiety disorders: Results of the international consortium in psychiatric epidemiology. *Addictive Behaviors*, *23*, 893-907.
Michell, J. (2005). The logic of measurement: A realist overview. *Measurement*, *38*, 285-294.

Middeldorp, C. M., Cath, D. C., Beem, A. L., Willemsen, G., & Boomsma, D. I. (2008). Life events, anxious depression and personality: a prospective and genetic study. *Psychological Medicine*, *38*, 1557-1565.

Milgram, S. (1967). The small world problem. *Psychology Today*, *2*, 60-70.

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology*, *49*, 377-412.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, *80*, 252-283.

Mischel, W., & Shoda, Y. (1994). Personality psychology has two goals: Must it be two fields? *Psychological Inquiry*, *5*, 156-158.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*, 246-286.

Mischel, W., & Shoda, Y. (1998). Reconciling processing dynamics and personality dispositions. *Annual Review of Psychology*, *49*, 229-258.

Mitchell, K., & Porteus, D. (2009). Gwas for psychiatric disease: is the framework built on a solid foundation? *Genomics and Human Genetics*, *3*, 740-741.

Moffitt, T. E., Harrington, H., Caspi, A., Kim-Cohen, J., Goldberg, D., Gregory, A. M., & Poulton, R. (2007). Depression and generalized anxiety disorder: Cumulative and sequential comorbidity in a birth cohort followed prospectively to age 32 years. *Archives of General Psychiatry*, *64*, 651-660.

Molenaar, P. C. M. (2003). *State space techniques in structural equation modeling: Transformation of latent variables in and out of latent variable models.*

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientic psychology, this time forever. *Measurement*, *2*, 201-218.

Mongrain, M. (1993). Dependency and self-criticism located within the five-factor model of personality. *Personality and Individual Differences*, *15*, 455-462.

Monroe, S. M., Harkness, K., Simons, A. D., & Thase, M. E. (2001). Life stress and the symptoms of major depression. *The Journal of Nervous and Mental Disease*, *189*, 168-175.

Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychological Bulletin*, *110*, 406-425.

Monroe, S. M., Torres, L. D., Guillaumot, J., Harkness, K. L., Roberts, J. E., Frank, E., & Kupfer, D. (2006). Life stress and the long-term treatment course of recurrent depression: Iii. nonsevere life events predict recurrence for medicated patients over 3 years. *Journal of Consulting and Clinical Psychology*, *74*, 112-120.

Munafò, M. R., Durrant, C., Lewis, G., & Flint, J. (2009). Gene x environment interactions at the serotonin transporter locus. *Biological Psychiatry*, *65*, 211-219.

Murray, O., & Lopez, A. D. (1997). Global mortality, disability, and the contribution of risk factors. global burden of disease study. *The Lancet*, *349*, 1436-1442.

Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide.*

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os,

J. (2009). Experience sampling research in psychopathology: opening the black box of daily life. *Psychological Medicine*, *39*, 1533-1547.

Nakamura, M., Ueno, S., Sano, A., & Tanabe, H. (2000). The human serotonin transporter gene linked polymorphism (5-httlpr) shows ten novel allelic variants. *Molecular Psychiatry*, *5*, 32-38.

Nash, M. W., Huezo-Diaz, P., Williamson, R. J., Sterne, A., Purcell, S., Hoda, F., ... Sham, P. C. (2004). Genome-wide linkage analysis of a composite index of neuroticism and mood-related scales in extreme selected sibships. *Human Molecular Genetics*, *13*, 2173-2182.

Neale, M. C., & Kendler, K. S. (1995). Models of comorbidity for multifactorial disorders. *American Journal of Human Genetics*, *57*, 935-953.

Newman, M. E. J. (2001a). Scientific collaboration networks. ii. shortest path, weighted networks, and centrality. *Physical Review E Nonlinear Soft Matter Physics*, *64*, 016132.

Newman, M. E. J. (2001b). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences USA*, *98*, 404-409.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*, 8577-8582.

Newman, M. E. J. (2010). *Networks: An introduction*. Oxford University Press.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*, 026113.

Nierenberg, A. A., Ostacher, M. J., Huffman, J. C., Ametrano, R. M., Fava, M., & Perlis, R. H. (2008). A brief review of antidepressant efficacy, effectiveness, indications and usage for major depressive disorder. *Journal of Occupational and Environmental Medicine*, *50*, 428-436.

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology*, *109*, 504-511.

Ogilvie, A. D., Battersby, S., Fink, G., Harmar, A. J., Goodwin, G. M., Bubb, V. J., & Dale Smith, C. A. (1996). Polymorphism in serotonin transporter gene associated with susceptibility to major depression. *The Lancet*, *347*, 731-733.

Olsen, L. R., Mortensen, E. L., & Bech, P. (2004). Prevalence of major depression and stress indicators in the danish general population. *Acta Psychiatrica Scandinavica*, *109*, 96-103.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree and shortest paths. *Social Networks*, *32*, 245-251.

Oquendo, M. A., Barrere, A., Ellis, S. P., Li, S., Burke, A. K., Grunebaum, M., ... Mann, J. J. (2004). Instability of symptoms in recurrent major depression: a prospective study. *American Journal of Psychiatry*, *161*, 255-261.

Ownsworth, T. L., & Oei, T. P. S. (1998). Depression after traumatic brain injury: conceptualization and treatment considerations. *Brain Injury*, *12*, 735-751.

Paik, K., & Kumar, P. (2008). Emergence of self-similar tree network organization. *Complexity*, *13*, 30-37.

Palmer, T. N. (2001). A nonlinear dynamical perspective on model error: a proposal for non-local stochastic-dynamic parametrization in weather and

climate prediction models. *Quarterly Journal of the Royal Metereological Society*, *127*, 279-304.

Papageorgiou, C., & Wells, A. (2001). Metacognitive beliefs about rumination in recurrent major depression. *Cognitive and Behavioral Practice*, *8*, 160-164.

Patrick, C. J., Edens, J. F., Poythress, N. G., Lilienfeld, S. O., & Benning, S. D. (2006). Construct validity of the psychopathic personality inventory two-factor model with offenders. *Psychological Assessment*, *18*, 204-208.

Patzold, L. M., Richdale, A. L., & Tonge, B. J. (1998). An investigation into sleep characteristics of children with autism and asperger's disorder. *Journal of Pediatric Child Health*, *34*, 528-533.

Paykel, E. S. (2008). Partial remission, residual symptoms, and relapse in depression. *Dialogues in Clinical Neuroscience*, *10*, 431-437.

Paykel, E. S., Scott, J., Teasdale, J. D., Johnson, A. L., Garland, A., Moore, R., ... Pope, M. (1999). Prevention of relapse in residual depression by cognitive therapy. *Archives of General Therapy*, *56*, 829-835.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Penninx, B. W., Nolen, W. A., Lamers, F., Zitman, F. G., Smit, J. H., Spinhoven, P., ... Beekman, A. T. F. (2011). Two-year course of depressive and anxiety disorders: Results from the netherlands study of depression and anxiety (nesda). *Journal of Affective Disorders*, *133*, 76-85.

Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry*, *5*, 103-113.

Pisarchik, A. N., & Goswami, B. K. (2000). Annihilation of one of the coexisting attractors in a bistable system. *Physical Review Letters*, *84*, 1423-1426.

Plante, D. T., & Winkelman, J. W. (2008). Sleep disturbance in bipolar disorder: Therapeutic implications. *American Journal of Psychiatry*, *165*, 830-843.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioral genetics*. Worth Publishers.

Posternak, M. A., & Miller, I. (2001). Untreated short-term course of major depression: a meta-analysis of outcomes from studies using wait-list control groups. *Journal of Affective Disorders*, *66*, 139-146.

Posternak, M. A., Solomon, D. A., Leon, A. C., Mueller, T. I., Shea, M. T., Endicott, J., & Keller, M. B. (2006). The naturalistic course of unipolar major depression in the absence of somatic therapy. *Journal of Nervous and Mental Disease*, *194*, 324-329.

Prescott, C. A., Aggen, S. H., & Kendler, K. S. (2000). Sex-specific genetic influences on the comorbidity of alcoholism and major depression in a population-based sample of us twins. *Archives of General Psychiatry*, *57*, 803-811.

Quitkin, F. M., McGrath, P. J., Stewart, J. W., Taylor, B. P., & Klein, D. F. (1996). Can the effects of antidepressants can be observed in the first two weeks of treatment? *Neuropsychopharmacology*, *15*, 390-394.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Paedagogiske Institut.

Ravikumar, P., Wainwright, M. J., & Lafferty, J. D. (2010). High-dimensional ising model selection using l1-regularized logistic regression. *The Annals of Statistics*, *38*, 1287-1319.

Read, S. J., Monroe, B. M., Brownstein, A. L., Yang, Y., Chopra, G., & Miller, L. C. (2010). A neural network model of the structure and dynamics of human personality. *Psychological Review*, *117*, 61-92.

Reichardt, J., & Bornholdt, S. (2006). When are networks truly modular? *Physical Review*, *74*, 016110.

Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology*, *5*, 27-48.

Richters, J. E., & Hinshaw, S. P. (1999). The abduction of disorder in psychiatry. *Journal of Abnormal Psychology*, *108*(438-445).

Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: a study of twins reared together using the self- and peer report neo-ffi scales. *Journal of Personality*, *65*, 449-475.

Rijsdijk, F. V., Sham, P. C., Sterne, A., Purcell, S., McGuffin, P., Farmer, A., ... Plomin, R. (2001). Life events and depressoin in a community sample of siblings. *Psychological Medicine*, *31*, 401-410.

Robins, C. J., & Block, P. (1988). Personal vulnerability, life events, and depressive symptoms: a test of a specific interactional model. *Journal of Personality and Social Psychology*, *54*, 847-852.

Roelofs, J., Huibers, M., Peeters, F., & Arntz, A. (2008). Effects of neuroticism on depression and anxiety: Rumination as possible mediator. *Personality and Individual Differences*, *44*, 576-586.

Roelofs, J., Huibers, M., Peeters, F., Arntz, A., & van Os, J. (2008). Rumination and worrying as possible mediators in the relation between neuroticism and symptoms of depression and anxiety in clinically depressed individuals. *Behaviour Research and Therapy*, *46*, 1283-1289.

Rothbaum, B. O., & Schwartz, A. C. (2002). Exposure therapy for posttraumatic stress disorder. *American Journal of Psychotherapy*, *56*, 59-75.

Rubinov, M., Knock, S. A., Stam, C. J., Micheloyannis, S., Harris, A. W. F., Williams, L. M., & Breakspear, M. (2009). Small-world properties of nonlinear brain activity in schizophrenia. *Human Brain Mapping*, *30*, 403-416.

Ruch, G. M. (1924). *The improvement of the written examination*. Scott Foresman.

Ruscio, J. (2009). A taxometric investigation of dsm-iv major depression in a large outpatient sample. interpretable structural results depend on the mode of assessment. *Assessment*, *16*, 127-144.

Ruscio, J., Zimmerman, M., McGlinchey, J. B., Chelminski, I., & Young, D. (2007). Diagnosing major depressive disorder xi: A taxometric investigation of the structure underlying dsm-iv symptoms. *Journal of Nervous and Mental Disease*, *195*, 10-19.

Savage, M. (2005). *Liberalism is a mental disorder: Savage solutions*. Nelson Current.

Savla, J., Davey, A., Costa, P. T., & Whitfield, K. E. (2007). Replicating the neo-pi-r factor structure in african-american older adults. *Personality and Individual Differences*, *43*, 1279-1288.

Schacter, D. L., & Addis, D. R. (2009). On the nature of medial temporal lobe contributions to the constructive simulation of future events. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*,

364, 1245-1253.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., ... Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, *461*, 53-59.

Scheines, R., Spirtes, P., Glymour, C., Meek, C., & Richardson, T. (1996). *Tetrad 3: Tools for causal modeling - user's manual.*

Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, *31*, 45-53.

Schoevers, R. A., Deeg, D. J. H., van Tilburg, W., & Beekman, A. T. F. (2005). Anxiety and depression: Co-occurrence and longitudinal patterns in elderly patients. *American Journal of Geriatric Psychiatry*, *13*, 31-39.

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse.* Guilford Press.

Seligman, L. D., & Ollendick, T. H. (1998). Comorbidity of anxiety and depression in children and adolescents: An integrative review. *Clinical Child and Family Psychology Review*, *1*, 125-144.

Serretti, A., Kato, M., De Ronchi, D., & Kinoshita, T. (2007). Meta-analysis of serotonin transporter gene promoter polymorphism (5-httlpr) association with selective serotonin reuptake inhibitor efficacy in depressed patients. *Molecular Psychiatry*, *12*, 247-257.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*, 2498-2504.

Sharot, R., Riccardi, A. M., Raio, C. M., & Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature*, *450*, 102-105.

Sheline, Y. I., Wang, P. W., Gado, M. H., Csernansky, J. G., & Vannier, M. W. (1996). Hippocampal atrophy in recurrent major depression. *Proceedings of the National Academy of Sciences USA*, *93*, 3908-3913.

Shoda, Y., LeeTiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and distinctiveness from intra- and interpersonal interactions. *Personality and Social Psychology Review*, *6*, 316-325.

Skinner, B. F. (1987). Whatever happened to psychology as the science of behavior? *American Psychologist*, *42*, 780-786.

Sklar, P. (2002). Linkage analysis in psychiatric disorders: the emerging picture. *Annual Review of Genomics and Human Genetics*, *3*, 371-413.

Slade, T. (2007). Taxometric investigation of depression: Evidence of consistent latent structure across clinical and community samples. *Australian and New Zealand Journal of Psychiatry*, *41*, 403-410.

Slade, T., & Andrews, G. (2005). Latent structure of depression in a community sample: a taxometric analysis. *Psychological Medicine*, *35*, 489-497.

Slavich, G. M., Thornton, T., Torres, L. D., Monroe, S. M., & Gotlib, I. H. (2009). Targeted rejection predicts hastened onset of major depression. *Journal of Social and Clinical Psychology*, *28*, 223-243.

Solé, R. V., Ferrer-Cancho, R., Montoya, J. M., & Valverde, S. (2000). Selection, tinkering, and emergence in complex networks. *Complexity*, *8*, 20-33.

Solomon, A., Haaga, D. A. F., & Arnow, B. A. (2001). Is clinical depression distinct from subthreshold depressive symptoms? a review of the continuity issue in depression research. *The Journal of Nervous and Mental Disease*, *189*, 498-506.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search*. MIT Press.

Spirtes, P., & Glymour, C. N. (1991). An algorithm for fast recovery of sparse causal graphs. *Social Science Computer Review*, *9*, 62-72.

Spitzer, R. L. (1973). A proposal about homosexuality and the apa nomenclature: Homosexuality as an irregular form of sexual behavior and sexual orientation disturbance as a psychiatric disorder. a symposium: should homosexuality be in the apa nomenclature? *American Journal of Psychiatry*, *130*, 1207-1216.

Spitzer, R. L. (1999). Harmful dysfunction and the dsm definition of mental disorder. *Journal of Abnormal Psychology*, *108*, 430-432.

Spitzer, R. L., & Endicott, J. (1978). Medical and mental disorder: Proposed definition and criteria. In D. F. Klein & R. L. Spitzer (Eds.), *Critical issues in psychiatric diagnosis*. Raven Press.

Stagner, R. (1951). Homeostasis as a unifying concept in personality theory. *Psychological Review*, *58*, 5-17.

Stefanis, N. (2008). Genes do not read dsm-iv: Implications for psychosis classification. *Annals of General Psychiatry*, *7*, S68.

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, *13*, 389-408.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, *410*, 268-276.

Sullivan, P. (2011). Letter to the editor: don't give up on gwas. *Molecular Psychiatry*, *DOI: 10.1038/mp.2011.94*.

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: review and meta-analysis. *American Journal of Psychiatry*, *157*, 1552-1562.

Tellegen, A. (1991). Personality traits: issues of definition, evidence and assessment. In D. Cicchetti & W. Grove (Eds.), *Thinking clearly about psychology: essays in honor of paul everett meehl* (p. 10-35). University of Minnesota Press.

Terracciano, A., Costa, P. T., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin*, *32*, 999-1009.

Terracciano, A., Lockenhoff, C. E., Crum, R. M., Bienvenu, O. J., & Costa, P. T. (2008). Five-factor model personality traits profiles of drug users. *BMC Psychiatry*, *8*, 22.

Terracciano, A., Sanna, S., Uda, M., Deiana, B., Usala, G., Busonero, F., ... Costa, P. T. (2010). Genome-wide association scan for five major dimensions of personality. *Molecular Psychiatry*, *15*, 647-656.

Teschl, G. (2008). *Ordinary differential equations and dynamical systems*. Gerald Teschl.

Thom, R. (1975). *Structural stability and morphogenesis*. Benjamin Press.

Thornton, T. (2000). Mental illness and reductionism: Can functions be naturalized? *Philosophy, Psychiatry, and Psychology*, *7*, 67-76.

Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, *41*, 1-32.

Trivedi, M. H., Rush, A. J., Wisniewski, S. R., Nierenberg, A. A., Warden, D., Ritz, L., ... Fava, M. (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in star*d: implications for clinical practice. *American Journal of Psychiatry*, *163*, 28-40.

Tryon, R. C. (1935). A theory of psychological components: An alternative to "mathematical factors". *Psychological Review*, *42*, 425-454.

Turkheimer, E. (1998). Heritability and biological explanation. *Psychological Review*, *105*, 782-791.

van Borkulo, C. D., Borsboom, D., Nivard, M. G., & Cramer, A. O. J. (2011). *Netlogo symptom spread model* (Tech. Rep.). http://ccl.northwestern.edu/netlogo/models/community/SymptomIL: Northwestern University Center for Connected Learning and Computer-Based Modeling: .

van den Brink, W., Hendriks, V. M., & van Ree, J. M. (1999). Medical coprescription of heroin to chronic, treatment-resistant methadone patients in the netherlands. *Journal of Drug Issues*, *29*, 587-608.

van der Maas, H. L. J., Dolan, C., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*, 842-861.

van der Maas, H. L. J., Molenaar, D., Maris, G., Kievit, R. A., & Borsboom, D. (2011). Cognitive psychology meets psychometric theory: on the relation between process models for decision making and latent variable models for individual differences. *Psychological Review*, *118*, 339-356.

van der Sluis, S., Kan, K. J., & Dolan, C. V. (2010). Consequences of a network view for genetic association studies. *Behavioral and Brain Sciences*, *33*, 173-174.

van der Sluis, S., Verhage, M., Posthuma, D., & Dolan, C. V. (2010). Phenotypic complexity, measurement bias, and poor phenotypic resolution contribute to the missing heritability problem in genetic association studies. *PLoS ONE*, *5*, 1-13.

van Geert, P. (1998). A dynamical systems model of basic developmental mechanisms: Piaget, vygotsky, and beyond. *Psychological Review*, *105*, 634-677.

van Nes, E. H., & Scheffer, M. (2007). Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *American Naturalist*, *169*, 738-747.

van Os, J., Hanssen, M., Bijl, R. V., & Vollebergh, W. (2001). Prevalence of psychotic disorders and community level of psychotic symptoms: an urban-rural comparison. *Archives of General Psychiatry*, *58*, 663-668.

van Os, J., & Jones, P. B. (2001). Neuroticism as a risk factor for schizophrenia. *Psychological Medicine*, *31*, 1129-1134.

Volkow, N. D., Wang, G. J., Fowler, J. S., Tomasi, D., & Telang, F. (2011). Addiction: Beyond dopamine reward circuitry. *Proceedings of the National Academy of Sciences USA*, *108*, 15037-15042.

Wadsworth, M. E., Hudziak, J. J., Heath, A. C., & Achenbach, T. M. (2001). Latent class analysis of child behavior checklist anxiety/depression in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 106-114.

Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist*, *47*, 373-388.

Wakefield, J. C. (1997). Diagnosing dsm-iv: Part i. dsm-iv and the concept of disorder. *Behaviour Research and Therapy*, *35*, 633-649.

Wakefield, J. C. (1999a). Evolutionary versus prototype analyses of the concept of disorder. *Journal of Abnormal Psychology*, *108*, 374-399.

Wakefield, J. C. (1999b). Mental disorder as a black box essentialist concept. *Journal of Abnormal Psychology*, *108*, 465-472.

Walker, F. O. (2007). Huntington's disease. *The Lancet*, *369*, 218-228.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxonometric procedures*. Thousand Oaks: Sage.

Wang, J., Fick, G., Adair, C., & Lai, D. (2007). Gender specific correlates of stigma toward depression in a canadian general population sample. *Journal of Affective Disorders*, *103*, 91-97.

Wang, X. F. (2002). Complex networks: Topology, dynamics and synchronization. *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, *12*, 885-916.

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for dsm-v. *Journal of Abnormal Psychology*, *114*, 522-536.

Watson, D., Weber, K., Assenheimer, J. S., Clark, L. A., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: I. evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology*, *104*, 3-14.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*, 440-442.

Wells, A., & Carter, K. (2001). Further tests of a cognitive model of generalized anxiety disorder: Metacognitions and worry in gad, panic disorder, social phobia, depression, and nonpatients. *Behavior Therapy*, *32*, 85-102.

Whiteford, H. A., Harris, M. G., McKeon, G., Baxter, A., Pennell, C., Barendregt, J. J., & Wang, J. (2012). Estimating remission from untreated major depression: a systematic review and meta-analysis. *Psychological Medicine*.

WHO. (n.d.). The global burden of disease report: 2004 update [Computer software manual].

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? a question for the diagnostic and statistical manual of mental disorders - fifth edition. *Journal of Abnormal Psychology*, *114*, 494-504.

Wiggins, J. S. (1984). Cattell's system from the perspective of mainstream personality theory. *Multivariate Behavior Research*, *19*, 176-190.

Wild, B., Eichler, M., Feiler, S., Friederich, H. J., Hartmann, M., Herzog, W., & Zipfel, S. (2007). Dynamic analysis of electronic diary data of obese patients with and without binge eating disorder. *Psychotherapy and Psychosomatics*, *76*, 250-252.

Wilensky, U. (1999). *Netlogo* (Tech. Rep.). Evanston, IL: Northwestern University Center for Connected Learning and Computer-Based Modeling.

Wirz-Justice, A., & van den Hoofdakker, R. (1999). Sleep deprivation in depression: what do we know, where do we go? *Biological Psychiatry*, *46*, 445-453.

Wong, M. L., Dong, C., Andreev, V., Arcos-Burgos, M., & Licinio, J. (2012). Prediction of susceptibiltiy to major depression by a model of interactions of multiple functional genetic variants and environmental factors. *Molecular Psychiatry*, *17*, 624-633.

Wu, J., & Bunney, W. (1990). The biological basis of an antidepressant response to sleep deprivation and relapse: review and hypothesis. *American Journal of Psychiatry*, *147*, 14-21.

Wurtzel, E. (2002). *Prozac nation: Young and depressed in america.* Riverhead Books.

Zachar, P. (2000). Psychiatric disorders are not natural kinds. *Philosophy, Psychiatry, and Psychology*, *7.3*, 167-182.

Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry*, *164*, 557-565.

Zarate, C. A., Singh, J. B., Carlson, P. J., Brutsche, N. E., Ameli, R., Luckenbaugh, D. A., ... Husseini, K. M. (2006). A randomized trial of an n-methyl-d-aspartate antagonist in treatment-resistant major depression. *Archives of General Psychiatry*, *63*, 856-864.

Zeeman, E. C. (1977). *Catastrophe theory: Selected papers.* Addison-Wesley.

Nederlandse Samenvatting

Wat is een psychische stoornis? Als we de populaire media en sommige stromingen in de gedragswetenschappen moeten geloven, dan is het antwoord op deze vraag vrij simpel. Psychische stoornissen zijn het beste te vergelijken met medische ziekten. Neem longkanker, een pathologische toestand in iemands lichaam die een set waarneembare symptomen - in het geval van longkanker bijvoorbeeld gewichtsverlies en een hardnekkige hoestprikkel - veroorzaakt. Zo wordt soms ook wel gedacht over bijvoorbeeld depressie: een pathologische toestand in iemands brein die een set waarneembare symptomen veroorzaakt zoals een sombere stemming, niet in slaap kunnen vallen 's nachts en vermoeid zijn. Dit klinkt aannemelijk, of toch niet?

Aan dit proefschrift ligt de opvatting ten grondslag dat het *niet* aannemelijk is dat de symptomen van een psychische stoornis als depressie op ongeveer dezelfde manier ontstaan als de symptomen van longkanker; of van het syndroom van Down. Voor deze stelling zijn twee hoofdredenen aan te voeren. De eerste betreft het feit dat in het geval van de longtumor en het syndroom van Down het volstrekt helder is wat de pathologische toestand is die de symptomen van deze ziekten veroorzaakt: in het geval van longkanker is dat een kwaadaardige tumor in de longen, in het geval van het syndroom van Down is dat een extra (gedeeltelijke) kopie van chromosoom 21. Voor psychische stoornissen is dit niet het geval: er is geen enkele pathologische toestand waarvan we, bijvoorbeeld, met zekerheid kunnen stellen dat dat *de* oorzaak is van depressieve symptomen. Zeker, er zijn diverse pathologische toestanden in het brein van (sommige) depressieve mensen aangetroffen, bijvoorbeeld een tekort aan serotonine (antidepressiva vullen dit tekort aan); echter, deze pathologie wordt bijvoorbeeld ook regelmatig aangetroffen bij patiënten met andere psychische stoornissen zoals verslaving en angststoornissen. In het geval van medische ziekten zou dit betekenen dat bijvoorbeeld een extra (gedeeltelijke) kopie van chromosoom 21 zowel het syndroom van Down zou kunnen veroorzaken als chronische verkoudheid. Juist het feit dat die extra chromosomale kopie een uniek effect sorteert - namelijk: het syndroom van Down - maakt dat we dat een afgebakende ziekte kunnen noemen, met een unieke ontstaansgeschiedenis en een unieke set symptomen. Dit is voor psychische stoornissen zeker niet het geval. Een tweede belangrijke reden om te twijfelen aan de aannemelijkheid van het medische model voor psychische stoornissen is dat in het geval van veel medische ziekten zoals longkanker geen *directe* verbanden bestaan tussen de symptomen zelf: gewichtsverlies (symptoom 1) veroorzaakt geen hardnekkige hoestprikkel (symptoom 2), of andersom. Het is de kwaadaardige tumor in de longen waardoor een patiënt gewicht verliest en last heeft van een hardnekkige hoestprikkel. In dit proefschrift betoog ik het tegenovergestelde: symptomen van psychische stoornissen hebben juist wel directe interacties met elkaar. Sterker nog, het zijn juist deze interacties die maken dat iemand een stoornis kan krijgen. Neem bijvoorbeeld niet slapen en vermoeid zijn, twee symptomen van depressie. Hoe zou het komen dat niet slapen en vermoeid zijn vaak samen ervaren worden door mensen? Doordat, zoals het medische

model voorschrijft, de stoornis 'depressie' deze klachten veroorzaakt? Of, omdat, wat in dit proefschrift betoogd wordt, moe zijn de *consequentie* is van niet slapen?

In de hoofdstukken 2 en 3 zet ik deze nieuwe *netwerkbenadering* van psychische stoornissen uiteen. Ieder persoon heeft zijn/haar eigen netwerk van symptomen en het is aannemelijk dat deze persoonsnetwerken van elkaar verschillen. Zo kan het bijvoorbeeld zijn dat bij de ene persoon een depressieve stemming sterk verbonden is met slecht in slaap kunnen komen 's nachts (depressieve stemming → slaapproblemen) maar dat dat voor een andere persoon helemaal niet geldt. Een depressieve stoornis ontwikkelen in een netwerkbenadering is dan geen kwestie van, zoals in het medische model, een abnormaliteit in het brein die bij iedereen op dezelfde manier symptomen veroorzaakt; maar een kwestie van symptomen die elkaar blijven 'aansteken' door directe interacties en er zo voor zorgen dat de persoon in kwestie in een vicieuze cirkel terecht kan komen die niet zonder hulp te doorbreken valt: bijvoorbeeld een persoon die eerst slecht slaapt waardoor vermoeidheid en concentratieproblemen op het werk ontstaan; als gevolg waarvan deze persoon zich schuldig voelt over het onderpresteren op het werk en daardoor een sombere stemming ontwikkelt waardoor die persoon vervolgens 's nachts wakker ligt van het piekeren. Zo ontstaat een kettingreactie of vicieuze cirkel, slaapproblemen → vermoeidheid en concentratieproblemen → schuldgevoelens → depressieve stemming → slaapproblemen, die zichzelf in stand houdt doordat aan het einde van de ketting de slaapproblemen weer het begin van een nieuwe kettingreactie vormen. Op verschillende plekken in het proefschrift wordt betoogd dat het aanhangen van een netwerkbenadering van psychische stoornissen mogelijk omvangrijke implicaties kan hebben voor hoe mensen in de klinische praktijk gediagnosticeerd en behandeld worden. Bijvoorbeeld, als er geen pathologische toestand, een gemeenschappelijke oorzaak, van depressiesymptomen is, dan heeft het ook weinig zin om deze te behandelen. Vanuit een netwerkbenadering zou het vooral zinnig zijn om individuele symptomen en de relaties tussen deze symptomen te behandelen.

Evidentie die consistent is met een netwerkbenadering van psychische stoornissen wordt behandeld in de hoofdstukken 4 en 5. In Hoofdstuk 4 laat ik zien dat verschillende stressvolle gebeurtenissen, zoals bijvoorbeeld een beëindigde liefdesrelatie, een verschillende invloed hebben op de individuele symptomen van depressie. Zo leiden gezondheidsproblemen tot meer suïcidale gedachten en een somberdere stemming in vergelijking met bijvoorbeeld een beëindigde liefdesrelatie. Dit is in overeenstemming met de netwerkbenadering maar strijdig met het medische model: immers, als de symptomen van depressie een gemeenschappelijke oorzaak zouden hebben (bijvoorbeeld een serotoninetekort) dan zouden stressvolle gebeurtenissen invloed moeten hebben op die gemeenschappelijke oorzaak, die dan vervolgens weer een set symptomen veroorzaakt. Andere evidentie in lijn met een netwerkbenadering van psychische stoornissen bespreek ik in Hoofdstuk 5 waarin ik laat zien dat depressiedata, gesimuleerd onder een netwerkmodel, allerlei kenmerken vertonen die karakteristiek zijn voor depressie. Het is bijvoorbeeld een bekend fenomeen dat een substantieel deel van de mensen met een depressie spontaan weer van deze depressie herstelt, dat wil zeggen zonder interventie van een therapeut of het gebruik van medicatie. Dit is precies wat ook gevonden wordt in de gesimuleerde data.

In de hoofdstukken 6 en 7 pleit ik ook voor een netwerkbenadering van persoonlijkheid. De Big Five is een welbekende karakterisering van de normale persoonlijkheid aan de hand van vijf trekken; extraversie, goedaardigheid, zorgvuldigheid, emotionele stabiliteit en openheid voor ervaringen. De theorie stelt dat iemand bijvoorbeeld extravert is en dat deze trek vervolgens allerlei waarneembaar gedrag veroorzaakt dat wij allemaal herkennen als zijnde extravert; bijvoorbeeld graag naar feestjes gaan en belangstelling hebben voor het ontmoeten van nieuwe mensen. Ik betoog dat ook deze trekken, net als psychische stoornissen, geen entiteiten zijn die ergens in iemands hoofd zitten, maar dat het hebben van een bepaalde trek, extravert, het gevolg is van directe interacties tussen

cognities, emoties en gedrag: iemand heeft belangstelling voor het ontmoeten van nieuwe mensen en gaat daardoor graag naar feestjes (nieuwe mensen ontmoeten → graag naar feestjes gaan).

Hoofdstuk 8 is een vrij praktisch hoofdstuk waarin diverse methoden worden uitgelegd en gedemonstreerd waarmee netwerken geconstrueerd en geanalyseerd kunnen worden. Een van deze methodes behelst het detecteren van causale relaties tussen symptomen; en ik laat met deze methode zien dat de comorbiditeit tussen depressie en gegeneraliseerde angst, dat wil zeggen het gelijktijdig te lijden hebben van deze twee stoornissen, voornamelijk verklaard kan worden door de symptomen die deze stoornissen met elkaar gemeenschappelijk hebben; zoals bijvoorbeeld slaapproblemen en vermoeidheid.

In dit proefschrift is een nieuwe benadering, de *netwerkbenadering*, voorgesteld die, zo wordt betoogd, meer recht doet aan de complexe realiteit van psychische stoornissen en persoonlijkheidstrekken. De mens is in zijn gedrag en gevoelens een complex wezen, en het is de opdracht van de gedragswetenschappen om deze complexiteit in kaart te brengen en, waar mogelijk, te begrijpen.

Dankwoord

Als iemand me aan het begin van mijn studie Psychologie had verteld dat ik ooit AiO zou worden bij Methodenleer, dan had ik die persoon hartelijk uitgelachen. Tijdens het grootste gedeelte van mijn studie wilde ik AiO worden bij Klinische Psychologie: het liefste wilde ik werken met forensische patiënten, en ik had daartoe al stage gelopen in de Van der Hoevenkliniek in Utrecht. Maar in de laatste fase van de research master raakte ik steeds meer geïnteresseerd in theoretische en methodologische vraagstukken. Precies op dat moment kreeg Denny een Vidi beurs voor een onderzoeksproject naar psychologische fenomenen als netwerken. Ik had stiekem nog nooit van Denny gehoord maar Eric-Jan - wiens onderwijsassistent ik op dat moment was - vond dat ik maar eens bij Denny moest binnenstappen en dirigeerde mij op een gegeven middag resoluut Denny's kamer in. Dankzij die welhaast letterlijke duw van EJ nam mijn beginnende leven in academia een totaal andere wending dan ik ooit voor mogelijk had gehouden; en het is de beste wending geweest die ik me maar had kunnen wensen.

Ik vind het moeilijk om de precieze woorden te vinden om je te bedanken Denny, want wat je voor me hebt betekend, gaat een stuk verder dan dat je mijn AiO tijd fijn begeleid hebt: je bent niet zozeer mijn begeleider geweest als wel mijn mentor, inspirator en grote voorbeeld. Je kan als geen ander iets opmerken of een vraag stellen, ogenschijnlijk terloops, waarvan ik de diepere betekenis of significantie pas begreep nadat ik er minimaal een week in totale vertwijfeling en ontreddering over had nagedacht. Dat soort opmerkingen en vragen hebben me zo enorm geïnspireerd om de grenzen van wat ik dacht dat ik kon op te zoeken; en deze grenzen een klein stukje te verleggen. En die grenzen kon ik ook verleggen omdat je me altijd de ruimte hebt gegeven om mijn eigen ideeën te ontwikkelen en me alle mogelijkheden hebt geboden om mijn methodologische kennis te verdiepen. Al zal ik nooit, echt nooit, een verwassen t-shirt van de Math Psych conferentie uit 1996 dragen, dankzij jou ben en voel ik me nu een methodoloog. Ik dank je ook zeer voor wat je me hebt geleerd op het menselijke vlak: niet alleen herinner ik me hoe meelevend en begripvol je was op momenten dat ik het moeilijk had, maar ook heb je me laten zien hoe belangrijk het is om de kwaliteiten van een ieder te herkennen en te waarderen. Ik zal het nauw samenwerken met jou enorm missen.

Als Han "nog een klein dingetje" heeft over een paper dat je van plan bent om op te sturen, dan kan je er gevoeglijk van uit gaan dat je minimaal een week bezig bent met het reviseren van dat paper. Ik dank je, Han, voor al die 'kleine' dingetjes die me altijd scherp hielden en me dwongen om een nog beter paper te schrijven. Ook ben ik je bijzonder dankbaar voor de manier waarop je leiding geeft aan onze vakgroep: je hebt oog voor ieders belangen, en je weet als geen ander hoe je mensen in stelling moet brengen voor verdere posities in academia. Ik heb mijn huidige baan voor een belangrijk deel aan jou te danken.

Ik dank alle overige leden van de promotiecommissie zeer voor het lezen en beoordelen van mijn proefschrift. Marten, ik koester mooie herinneringen aan de legendarische

'Orchard sessions'. EJ, ik ben heel blij dat je me bijna in Denny's kamer geduwd hebt in 2007, en ik denk met veel plezier terug aan al onze gezellige borrels en etentjes. Sophie, je maakt me ontzettend aan het lachen; dank voor alle gezellige theekransjes, ik hoop dat er nog vele gaan volgen.

Het was een voorrecht om te mogen werken met een keur aan interessante, inspirerende en gezellige collega's. In het bijzonder denk ik met een grote glimlach terug aan mijn oude 'roomies': Dylan (mijn Breezer-partner-in-crime en de gezelligste psychometricus die ik ken), Don ("That's gooood" en andere Peanut gekkigheid op de vrijdagmiddag, en de bedenker van een passende bijnaam voor ondergetekende), en Ruud (de minst luidruchtige van A5.02, staat altijd garant voor diepgaande discussies en flauwe grappen, en geneerde zich totaal niet om een koptelefoon op te zetten om te ontsnappen aan geklets van ondergetekende). Mariska, dank je wel dat je me geïnspireerd hebt om een goede docent te willen zijn; en ik mis je groenten, stijlvolle pumps, warmte en agenda op vier A4-tjes nog zeer regelmatig. Marjan, dank je voor je vriendschap: onze urenlange gesprekken over werk maar ook over onze kinderen zijn me zeer veel waard; en samen met Josine heb je me door de laatste zware maanden van mijn AiO tijd gesleept.

I owe a great deal to all the wonderful people that I met and worked with during my stay at VATSPUD in Richmond, Virginia. Dear Dawn, Shaunna, Brad, Liz, Ananda, Daniel, Mike, and many others: thank you for all your kindness, manicures, great evenings with loads of laughter and good food, cupcakes, champaign cocktails and dancing; thanks to you Richmond is one of the few places that I call home. Steve, thank you for the numerous times that you helped me out (from airport pick ups to fitting complex factor models) and for the many - sometimes pretty fierce - discussions about the nature of psychological phenomena. Ken, most of what I think I know about psychiatry, I know it because you taught me. I'm so thankful for the many hours we spent in your office during which you shared not only your knowledge and experience with me; but also your passion for the twin study, you reminded me of the significance of the individual people behind the numbers in my data file. I also have fond memories of the wonderful Thanksgiving dinner I shared with you and your family, and of the afternoon we spent in the movie theater during the French film festival. Thank you.

Ik heb fantastische vrienden: veel dank voor alle gesprekken, feestjes, uithuilschouders, kaufrausches, grappen en grollen. Over twee vrienden in het bijzonder wil ik nog iets meer zeggen, mijn paranimfen Eveline en Rogier. Eveline, je bent een van mijn allerliefste vriendinnetjes: je hebt mij door diepe dalen geholpen maar ook op de pieken ben je er altijd bij (ik zeg bijvoorbeeld Naughty by Night en laat het daar verder bij). Je maakt me vrolijk, je bent een prachtmens, van buiten en van binnen. Rogier, je bent een ontzettend grappige, slimme, gezellige maar bovenal buitengewoon goede vriend; ik denk bijvoorbeeld aan ons gezamenlijke tripje naar Washington waar jij mij, net zwanger en vreselijk misselijk, middels taxi's en tussenstops in winkels waar mannen niet van houden, door de stad heen hebt geloodst. Maar ook zal ik nooit vergeten hoe je me gesteund hebt toen ik aan huis was gekluisterd door mijn rug. Ik heb jou, Anne-Laura en Flynn voor altijd in mijn hart gesloten.

Lieve mama, jij hebt me geleerd om onbaatzuchtig te zijn, om het goede in mensen te zien en wat het betekent om je echt met hart en ziel ergens voor in te zetten. Ik was niet waar ik nu ben zonder jou. En ik geef toe dat je toch gelijk had: Wijsheid komt met de jaren. Ik hoop nog heel lang van je te mogen leren: je t'aime. Casper: liefde van mijn leven en mijn allerbeste vriend. Dank voor onze liefde, gezelligheid, en voor al die avonden dat je me mijn gang hebt laten gaan op mijn laptop; en voor alle momenten dat je me van deze zelfde laptop hebt weten los te weken. Het leven is een feest met jou. Liefste Isa Nora, het leven is zoveel mooier sinds ik door jouw bril mag meekijken naar de wereld; het is een voorrecht om jouw moeder te zijn.

Allerliefste papa, wat ben ik verdrietig omdat je nooit zal weten dat het me gelukt is, dat ik een proefschrift heb afgeleverd. Ik ben een atheïst in hart en nieren maar soms, op dit soort momenten, hoop ik vurig dat jij op een of andere manier mee kan kijken; zodat je kan zien dat ik, ondanks dat ik je nog iedere dag vreselijk mis, mijn dromen probeer te verwezenlijken. Ik draag dit proefschrift aan je op en dank je voor alles wat je me gegeven hebt. Aye aye, sir.

Angélique
Amsterdam, juli 2013