Comorbidity: A network perspective

Angélique O. J. Cramer  
University of Amsterdam  
Department of Psychology  
Roetersstraat 15  
1018 WB Amsterdam  
The Netherlands  
A.O.J.Cramer@uva.nl

Lourens J. Waldorp  
L.J.Waldorp@uva.nl

Han L. J. van der Maas  
H.L.J.vanderMaas@uva.nl

Denny Borsboom  
D.Borsboom@uva.nl

Author note: This work was supported by NWO innovational research grant no. 451-03-068.

Abstract: The pivotal problem of comorbidity research lies in the psychometric foundation it rests on, i.e., 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, 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.

Keywords: comorbidity, latent variable models, complex networks, major depression, generalized anxiety
Introduction

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., Albert, Rosso, Maina & Bogetto, 2008; Brown, Anthony & 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; Brown, Campbell, Lehman, Grisham & Mancill, 2001; Kashani et al., 1987; Kessler et al., 1994; Kessler et al., 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 largely an artifact of the diagnostic system.

In this paper, 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, Mellenbergh & Van Heerden, 2003; Borsboom, 2005, 2008). 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 such 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 paper, we provide an in depth discussion of these problems, and show that the network approach avoids them.
The structure of this paper is as follows. First, we introduce the network approach by contrasting it to the latent 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 (henceforth: MDD) and Generalized Anxiety Disorder (henceforth: GAD), based on data from the National Comorbidity Survey Replication\(^1\) (NCS-R; Kessler et al., 2004; Kessler, Berglund, Demler, Jin & Walters, 2005; Kessler, Chiu, Demler & Walters, 2005). 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?

2. 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/absence of certain observable variables (e.g., symptoms of MDD or party going behavior measuring extraversion; 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, 2008; Borsboom, Mellenbergh & Van Heerden, 2003, 2004; Hood, 2008): in this view, for instance, depression (i.e., the latent attribute) causes symptoms such as fatigue to occur.

\(^1\) 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 carried out 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 paper. The symptoms that participants reported within one disorder all occurred within the same time frame.
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 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 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 correlations or odds ratios between those proxies. Although this methodology has yielded important insights (e.g., Brown, Campbell, Lehman, Grisham & Mancill, 2001; Kessler et al., 1994; Kessler, Chiu, Demler & Walters, 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, 2008).

![Figure 1](image_url). A model of comorbidity between disorder A and disorder B, under the standard assumptions of latent variable modeling. The circles represent the disorders (i.e., latent variables) while 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 disorder A and disorder 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, Neale & Kendler, 2005; 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 (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 (e.g., see Hartman
et al., 2001). However, we do suggest that mental disorders may not explain covariation between symptoms in the way a latent variable model pictures the situation. If this is so, then even though the application of latent variable modeling may have considerable instrumentalistic utility (e.g., in facilitating predictions or gauging rough differences between people), one cannot plausibly 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 paper. 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 & Barabási, 1999, 2002; Boccaletti, Latora, Moreno, Chavez & Hwang, 2006; Strogatz, 2001; 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 due to a mental disorder as a common cause (Borsboom, 2008; 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 direct 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 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, such symptoms 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.
Figure 2. Comorbidity under a network approach. Disorder A consists of bidirectionally related symptoms $X_1$-$X_5$; disorder B consists of symptoms $Y_1$-$Y_5$. Symptoms $B_1$-$B_2$, are *bridge symptoms* that overlap between disorder A and B. In this model, comorbidity arises as a result of direct relations between the bridge symptoms of two 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 models with $n^{th}$ order interactions – that 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 that in Figure 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 only arises through connections between overlapping symptoms on the one hand and other symptoms of MDD/GAD on the other hand. We used the AIC to compare the

---

2 We did not collapse the six symptoms that overlap between MDD and GAD into three bridging 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 (for instance because the respondent did not experience chronic anxiety).
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 this term to mean that the relation between symptoms is real, i.e., is not spurious in the sense that a latent variable model assumes it to be. This does not imply, however, there may be no intermediate processes or attributes involved. 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 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 no 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 respond to it with fear by repeatedly exposing a patient to the feared object; e.g., see 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 only beneficial for 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.
3. An integrative method to visualize symptom associations through graphical models

Many of the efforts 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 (De Berg, Cheong, Van 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 paper), aiSee (http://www.aisee.com), and igraph (Csárdi & Nepusz, 2006; used in this paper 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 3. The legend for Figures 4, 5 and 6.
Among a plethora of possibilities to define and visualize both nodes and edges (e.g., see Boccaletti, Latora, Moreno, Chavez & Hwang, 2006; Krichel & Bakkalbasi, 2006), we propose an integrative method that, in our view, optimally visualizes key aspects of comorbidity on a symptom level. Figure 4 presents such a comorbidity network for MDD and GAD (see Figure 3 for the complete legend): 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 log odds ratio between two symptoms (i.e., strength of the association; results available on www.aojcramer.com): the higher the log odds ratio, the darker blue the edge between symptoms (note that other options exist to define some measure of the strength of the association between two symptoms, for instance tetrachoric correlations). 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 (e.g., see 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 4). First, we propose that from left to right (i.e., the x-axis), non-overlapping symptoms of two disorders are placed on the extreme left and right while the overlapping symptoms are placed in the middle of the graph. As such, one can immediately see whether comorbidity between two disorders runs mostly through the

---

4 This network is based on the NCS-R questionnaire that mostly contains dichotomous items. However, some of the items were not (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 on www.aojcramer.com.

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 \( \frac{n_{11}n_{22}}{n_{12}n_{21}} \) (see Agresti, 2002). Since the odds ratio scales between zero and infinity, with a value of one 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 that 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 dataset.

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.
overlapping symptoms or (also) exists independently from them. Second, we propose that from top to bottom (i.e., 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).

Figure 4. A comorbidity network for MDD and GAD. Larger nodes represent more frequent symptoms; darker circumference, higher centrality; thicker edges, higher frequency of co-occurrence; darker edges, stronger associations. Only edges with a log odds ratio higher than (-) 0.60 are represented. Centrally positioned nodes (mConc, gConc, mSleep, gSleep, mFatig, gFatig, mRest and gRest) represent overlapping symptoms. Non-overlapping MD symptoms are displayed on the left of the figure, non-overlapping GAD symptoms on the right.

4. 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 47). 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, Demler & Walters, 2005). However, on a diagnosis level, only

---

7 We have checked the stability of the results depicted in this figure by randomly splitting the sample in two and run all analyses for both groups separately. Those separate analyses revealed the same results and thus, we consider the components of Figure 4 to be stable.
respondents who display a certain number of MDD or GAD symptoms with a certain duration qualify for a diagnosis. Additionally, due to a hierarchical exclusion rule, the GAD diagnosis will not be assigned if its symptoms occur exclusively within the course of MDD (Brown & Barlow, 1992; Brown, Campbell, Lehman, Grisham & Mancill, 2001; Clark, Watson & Reynolds, 1995; Mineka, Watson & Clark, 1998; Watson, 2005). Since MDD and GAD are highly comorbid (e.g., see Brown et al., 2001; Brown, Chorpita & Barlow, 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 they obtained diagnoses or not. 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 the 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 associations 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, Watson & Clark, 1998) and as such, raise the question 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 (i.e., few edges incident in those nodes). This may appear surprising since in clinical practice, duration is key in determining the presence/absence of a mental disorder. However, if we consider medical illnesses as an analogy, the finding is potentially less surprising: cancer will be diagnosed if a malignant tumor is present, and that diagnosis is

8 Respondents with depressed mood/loss of interest for at least three days for more than one hour per day progressed to the section about the other symptoms. For generalized anxiety, respondents with chronic anxiety for at least one month progressed to the section about the other symptoms. As such, the fact that duration is weakly associated with the other MDD and GAD symptoms cannot be explained by people only progressing through the remainder of the interview if they fulfilled the respective duration criteria (more than two weeks for MDD and more than six months for GAD).
independent of how long the tumor has been present. Thus, we could argue that, in a network approach, MDD is present whenever some symptoms are present without considering the duration of those symptoms. This is not to say that duration is not an important factor at all. Consider again medical illnesses where duration is important in determining the course of action and subsequently, the probability of full recovery: the longer a malignant tumor has had time to grow and possibly spread, the more difficult it will be to treat it. Duration could fulfill the same role in determining the best course of action for treating mental disorders.

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; Clark & Watson, 1991; 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; Roelofs, Huibers, Peeters, Arntz & Van Os, 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 symptoms 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 such variables are portrayed in latent variable models commonly used in data analysis.

5. 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 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 in fact 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 4 (i.e., colors of the nodes), one immediately sees that, indeed, depressed mood ($m_{Dep}$) and loss of interest ($mint$) are far more central in the network than psychomotor ($m_{Rest}$) and weight problems ($m_{Weight}$). 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 thus 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, individuals who do not meet diagnostic criteria for MDD 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 Alice is not depressed and Bob is? In other words, based on diagnostic cut-offs, we may fail to disentangle symptom specific effects, because such cut-offs do not take the centrality of symptoms into account.

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 symptoms'. According to a latent variable perspective, it is not only perfectly defensible to entertain such a definition, 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?

---

\(^9\) 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.
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 (e.g., see Hubert, 1974). First, let’s 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 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; e.g., see Boccaletti, Latora, Moreno, Chavez & Hwang, 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 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 non-trivial 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 non-existent 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 and 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 a network) than others. Figure 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.

6. 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; Haslam & Ernst, 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 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 21st chromosome, and everyone with Down syndrome possesses that defining feature while everyone without Down 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 get the diagnosis of 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 the ‘MDD’ category – because he/she suffers from five or more symptoms for more than two weeks – as well as the ‘no MDD’ category if he/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 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 a refinement of the conceptual and measurement apparatus used to study them (e.g., Bollen, 2002; Borsboom, 2008). However, in the absence of such refinements, the acceptance of the latent variable hypothesis depends at least partly on its explanatory virtues (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 during the history of psychiatry and clinical psychology (Haslam, 2000; Haslam & Ernst, 2002; Jablensky, 2007; Kendell, 1975; Klein, 1978; Krueger & Markon, 2006; Lilienfeld & Marino, 1999; Spitzer, 1973, 1999; Richters & Hinshaw, 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 that could be given to it.

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/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 spin glass 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, Watson & Clark, 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 has been noticed quite some time ago (e.g., see Kendell, 1975; 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 separate the relevant networks). One 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 actually 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, depression, as well as many other psychological variables, is determined by genetic factors (Kendler, Gardner, Neale, & Prescott, 2001; McGue, & Christensen, 2003; Boomsma, Busjahn, & Peltonen, 2002). 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., one or two percent 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 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 polygenic, 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 and colleagues (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 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 each other, 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, it does not appear to be more likely to quickly progress from a few non-overlapping GAD symptoms to overlapping GAD symptoms and from there to MDD symptoms. 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 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 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 to the study of temporal dynamics of symptoms to be essential.

7. 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 approached this problem via different angles and with respect to different disorders, and the majority 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 (e.g., see Biederman, Faraone, Mick & Lelon, 1995; Bleich, Koslowsky, Dolev & Lerer, 1997; Franklin & Zimmerman, 2001; Kessler, DuPont, Berglund & Wittchen, 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 and colleagues (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 while 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 justifiably 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.

This is not to say we think that the removal of overlapping symptoms to study its effects 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 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 investigated 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 4, resulting in Figure 5 (see Figure 3 for the legend). 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 worrying about (gEvent).

![Figure 5](image)

**Figure 5.** The comorbidity network for MDD and 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 4.

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). Figure 6 presents their comorbidity network without the overlapping symptoms (see Figure 3 for the legend). This figure leaves no room for doubt about the importance of overlapping

---

10 The contingency tables as well as the computational script (made in R) are available on [www.aojeramer.com](http://www.aojeramer.com). We have checked the stability of the results depicted in this figure by randomly splitting the sample in two and run all analyses for both groups separately. Those separate analyses revealed the same results and thus, we consider the components of Figure 6 to be stable.
symptoms: all symptoms are more frequent and co-occur more frequently, and having one symptom increases the odds of having another one substantially (and thus the node strength) compared to the comorbidity network in Figure 5. Taken all results together, it is likely that overlapping symptoms play a more important role in explaining comorbidity than was originally thought.

Figure 6. A comorbidity network for MDD and 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 5.

8. Conclusions and future directions

In this paper, 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 each other. 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 4) inequality of symptoms. Also, we have proposed an integrative method to visualize comorbidity networks based on bivariate associations.

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 due to 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: 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 to move towards one
attractor and remain there in a stable state (i.e., monostable systems; e.g., see 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., system moves towards 'depressed' attractor state), and go (i.e., system moves towards '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/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, etc. 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 would 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 (e.g., see Paik & Kumar; Solé, Ferrer-Cancho, Montoya & Valverde, 2000). An important additional question is how dynamical properties of complex systems relate statistically and conceptually to interindivdual 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 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.

9. References


