Kinds versus continua

A review of psychometric approaches to uncover the structure of psychiatric constructs

Borsboom, D.; Rhemtulla, M.; Cramer, A.O.J.; van der Maas, H.L.J.; Scheffer, M.; Dolan, C.V.

DOI
10.1017/S0033291715001944

Publication date
2016

Document Version
Final published version

Published in
Psychological Medicine

License
Article 25fa Dutch Copyright Act

Psychological Medicine (2016), **46**, 1567–1579. © Cambridge University Press 2016
doi:10.1017/S0033291715001944

REVIEW ARTICLE

Kinds *versus* continua: a review of psychometric approaches to uncover the structure of psychiatric constructs

D. Borsboom[1]*, M. Rhemtulla[1], A. O. J. Cramer[1], H. L. J. van der Maas[1], M. Scheffer[2] and C. V. Dolan[3]

[1] *Department of Psychology, University of Amsterdam, Weesperplein 4, Amsterdam 1018 XA, The Netherlands*
[2] *Department of Aquatic Ecology and Water Quality Management, Wageningen University, 6700 AA Wageningen, The Netherlands*
[3] *Department of Biological Psychology, VU University, 1081 BT Amsterdam, The Netherlands*

The question of whether psychopathology constructs are discrete kinds or continuous dimensions represents an important issue in clinical psychology and psychiatry. The present paper reviews psychometric modelling approaches that can be used to investigate this question through the application of statistical models. The relation between constructs and indicator variables in models with categorical and continuous latent variables is discussed, as are techniques specifically designed to address the distinction between latent categories as opposed to continua (taxometrics). In addition, we examine latent variable models that allow latent structures to have both continuous and categorical characteristics, such as factor mixture models and grade-of-membership models. Finally, we discuss recent alternative approaches based on network analysis and dynamical systems theory, which entail that the structure of constructs may be continuous for some individuals but categorical for others. Our evaluation of the psychometric literature shows that the kinds–continua distinction is considerably more subtle than is often presupposed in research; in particular, the hypotheses of kinds and continua are not mutually exclusive or exhaustive. We discuss opportunities to go beyond current research on the issue by using dynamical systems models, intra-individual time series and experimental manipulations.

Received 23 January 2015; Revised 3 September 2015; Accepted 3 September 2015; First published online 21 March 2016

Key words: Dynamical systems, latent variable models, network models, psychometrics, taximetrics.

Introduction

The question of whether mental disorders should be thought of as discrete categories or as continua represents an important issue in clinical psychology and psychiatry. The current setup of diagnostic systems such as the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5; American Psychiatric Association, 2013) and the 10th revision of the International Classification of Diseases (ICD-10; World Health Organization, 1992) typically adheres to a categorical model, in which discrete diagnoses are based on patterns of symptoms.

This approach is rooted in psychiatric traditions that go back to the work of Emil Kraepelin (e.g. see Kraepelin & Dierendorf, 1915), who laid the foundation for a psychiatric categorization system that views the science and diagnosis of mental disorders as a branch of medicine. In medicine, tracing observable symptoms (e.g. foggy eyesight, headaches) to specific diseases (e.g. a tumour in the brain) plays a central role (Hyland, 2011). In fact, the successes of modern medicine are predicated on the insight that, in many cases, treatment should be directed at diseases (e.g. removing the tumour) rather than, for instance, at the observable symptoms themselves, because in medicine diseases function as root causes (Borsboom & Cramer, 2013). In this scheme of thinking, the decision of which treatment to assign to an individual depends on which disease that person has, rather than on the observable symptoms. Thus, the task of the physician is to identify diseases through diagnosis, after which an appropriate treatment can be selected. This idea functions as a template for health care systems around the world, and mental health care is no exception.

The current health care model assumes that psychiatric categorizations 'carve nature at its joints', as Plato puts it. However, such categorizations often involve apparently arbitrary conventions. For instance, while the DSM-5 diagnosis of major depression requires five or more symptoms to be present, it is unclear whether the resulting categorization is empirically superior to one that would require four or six symptoms for a diagnosis. If such categorizations do not have parallels in reality, for instance because they are

* Address for correspondence: D. Borsboom, Department of Psychology, University of Amsterdam, Weesperplein 4, Amsterdam 1018 XA, The Netherlands.
(Email: d.borsboom@uva.nl)

essentially arbitrary cut-scores on a continuum, current practices may frustrate rather than facilitate scientific progress. For example, artificially dichotomizing a continuous variable can lead to dramatic drops in statistical power to detect effects on that variable (MacCallum *et al.* 2002; Van der Sluis *et al.* 2013). Also, such practices may lead to suboptimal treatment (or no treatment at all); both because evidence for diagnosis–treatment combinations is compromised, and because individuals seeking treatment may not fit the distinctions imposed by the system.

Thus, the question of whether psychopathology constructs are kinds or continua is of central importance to psychiatry and clinical psychology. In the past two decades, researchers have started to approach this question empirically (Meehl, 1992, 1995; Waller & Meehl, 1998). Collectively known as taxometrics, empirical approaches are based on the idea that multivariate distributions of psychometric indicators can be used to infer the structure of the underlying construct (Ruscio *et al.* 2006). This research programme has grown rapidly in recent years: Haslam *et al.*'s (2012) review discusses no fewer than 177 taxometric studies.

The present paper reviews such modelling approaches from a psychometric perspective. First, we discuss classic psychometric conceptualizations of latent variables and approaches to investigate them. Second, we examine psychometric models that integrate categorical and continuous features of latent constructs. Third, we review alternative approaches based on network modelling, which allow constructs to be continuous for some individuals but categorical for others. We conclude that current methodologies and data are insufficient to conclusively decide whether psychopathology constructs are continuous or categorical.

Measurement theoretical definitions of kinds and continua

The question of how to represent empirical phenomena (such as depression) using mathematical structures (such as continuous or categorical systems) is dealt with in measurement theory[1]† (e.g. Suppes & Zinnes, 1963; Krantz *et al.* 1971). Importantly, all measurement starts with categorization; namely with the formation of equivalence classes. Equivalence classes are sets of individuals who are exchangeable with respect to the attribute of interest (e.g. a psychiatric disorder). Thus, we can imagine that, according to some ideal method of observation, all depressed individuals would be assigned the same abstract symbol, say 'D', while the rest of the population would be assigned the symbol 'H'. If all of the individuals in the population can be assigned these labels unambiguously, and individuals with the same label are indistinguishable with respect to the attribute of interest (just like two electrons are indistinguishable with respect to their charge), then we end up with two categories, as in the leftmost panel of Fig. 1. This is a nominal measurement scale (Stevens, 1946), because even if the symbols may be numerals (e.g. 0 and 1), they do not carry numeric information but only serve to convey class membership.

Now, we may not succeed in finding an observational procedure (or even in agreeing on a hypothetical one) that in fact yields the desired equivalence classes. For instance, we may – and typically do – find that individuals who have been assigned the same label are not indistinguishable with respect to the attribute of interest: there appear to be significant differences between, say, cases of depression that feature the full range of symptoms and those that barely meet the diagnostic criteria. An intuitive way to accommodate this is to create more than two categories; e.g. 'no depression' (H), 'mild depression' (M) and 'severe depression' (S).

Because there are now three classes rather than two, next to the relation between individuals within classes (equivalence) we may also represent systematic relations between members of different classes. One way to do this is by invoking the concept of order, as represented by assigning numbers to the individuals in the classes (e.g. assign a 0 to all H-individuals, a 1 to all M-individuals, and a 2 to all S-individuals). Note that any numerical assignment that represents this order will do: the numbers still have no quantitative meaning (Michell, 1997, 1999). This representation, which is shown in the middle panel of Fig. 1, is known as an ordinal scale in measurement theory (Stevens, 1946) and as an ordered categorical structure in statistics (e.g. Agresti, 2013).

This procedure may, however, fail too. For example, we may find that even within the H, M and S-classes, there are non-trivial differences between individuals that we wish to represent. In this case we can break up these classes into even more subclasses; say M is split into M1, M2 and M3, and S is split into S1, S2 and S3. If (and only if) these additional subclasses also conform to order relations, such that $H < M1 < M2 < M3 < S1 < S2 < S3$, then we may represent them with a scale that starts to approach continuity: a continuous variable can be seen as an extension of this line of reasoning to infinitely many (possible) subclasses. The continuity hypothesis formally implies that (*a*) in between any two positions lies a third that can be empirically instantiated[2] (just like for any two people who are 1.8 and 1.9 m tall, we might find a third who is 1.85 m tall), and (*b*) that there are no

† The notes appear after the main text.

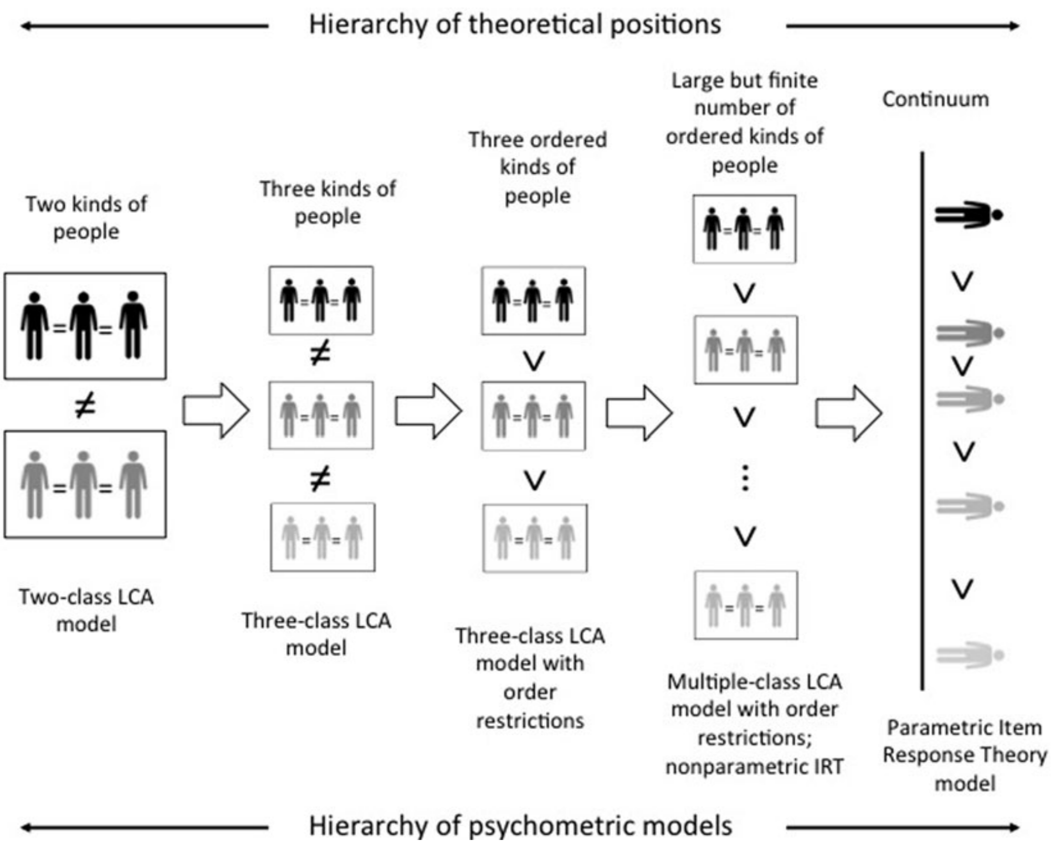

Fig. 1. How different theories on the structure of attributes (top) imply distinct latent variable models (bottom) in the case of dichotomous indicator variables. The top of the figure progresses from a theoretical structure with two kinds of people, via an ordered structure, to a continuum. The bottom of the figure shows the parallel psychometric progression from a latent class analysis (LCA) model (left) to a parametric item response theory (IRT) model (right) via ordered latent class models (or, equivalently, non-parametric IRT models).

gaps in the continuum (just like, within the normal spectrum, there are no impossible body heights). This situation is represented in the rightmost panel of Fig. 1.

In psychological terms, categorical representations line up naturally with an interpretation of disorders as discrete disease entities (e.g. tumours, which are present or not), while continuum hypotheses are most naturally consistent with the idea that a construct varies continuously in a population (e.g. bodily height, which everybody has in some quantity). In a continuous interpretation, the distinction between individuals with and without, say, a DSM-5 diagnosis would thus be analogous to the distinction between tall people and people of average height: it depends on the imposition of a cut-off score that does not reflect a gap that is inherent in the attribute itself. Note that this does not mean that the delineated categories are arbitrary in a general sense or are 'just' social conventions: the delineation of extremely tall people may not correspond to any clear cut-off, but surely extremely tall people possess qualities and problems that the rest of us do not have, and the concept of being extremely tall certainly does not rest primarily on a social convention.

Kinds and continua as psychometric entities

If mental disorders were directly observable, the task of categorizing them as continuous or categorical would be relatively straightforward, and could proceed exactly as illustrated before. In this case, we could actually create equivalence classes of individuals with the same disorder status, see how many of these suffice, and test whether they conform to order relations. All this is standard procedure in the natural sciences, where the basic equivalence relations (being equally heavy, being equally tall) can be determined directly through experiment (Trendler, 2009; Markus & Borsboom, 2012). With a sufficiently large number of neatly ordered categories (say, 30 levels of depression, which would differ exclusively in degree), few would object to treating depression as continuous, at least for most practical and research purposes[3].

Unfortunately, in psychology, we have no way to decide conclusively whether two individuals are 'equally depressed'. This means we cannot form the equivalence classes necessary for measurement theory to operate (Krantz *et al.* 1971; see also Trendler, 2009; Markus & Borsboom, 2013). The standard approach to dealing with this situation in psychology is to presume that, even though equivalence classes for theoretical entities like depression and anxiety are not subject to direct empirical determination, we may still entertain them as hypothetical entities purported to underlie the thoughts, feelings and behaviours that we do observe (Borsboom, 2005). Under this assumption, we may investigate these theoretical constructs indirectly, by conceptualizing them as the common cause of a set of symptoms or item responses (e.g. Cronbach & Meehl, 1955; Bollen, 1989; Reise & Waller, 2009).

For example, a psychometric model could specify the hypothesis that there are just two kinds of people (e.g. depressed and non-depressed individuals), and that these two kinds of people respond differently to the questions in a clinical interview; this would yield a formal representation of the hypothesis that disorders are discrete kinds (e.g. a latent class model; Lazarsfeld & Henry, 1968; McCutcheon, 1987). Alternatively, a model could hold that there are not just two discrete categories of individuals, but rather that individuals differ from each other in degree (e.g. everybody, including healthy individuals, has a certain level of depression). This would yield a formal representation of the hypothesis that depression is not a kind but a continuum[4]. If symptoms were recorded dichotomously, the resulting mathematical structure would be an item response theory (IRT) model (Reise & Waller, 2009). An IRT model states that the probability of endorsing an item depends is a monotonic function of a person's position on the measured latent variable, so that a higher position on that latent variable is associated with a higher probability of endorsing the item. The precise form of the item characteristic curve (ICC; which relates the item response probability to the latent variable) can vary across IRT models, but in standard applications is often taken to be a logistic curve.

Crossing these conceptualizations of latent structure with the structure of observations yields Bartholomew's (1987) classic taxonomy of measurement models as represented in Table 1.

Table 1. *Cross-classification of latent variable models for discrete* v. *continuous latent and observed variables*

	Latent variables	
	Discrete	Continuous
Observed variables		
Discrete	Latent class models	Item response theory models
Continuous	Latent profile models	Common factor models

The models in Table 1 are identical in that the latent and observed variables feature as independent and dependent variables, respectively (see also Mellenbergh, 1994). Also, in each of the latent variable models represented in Table 1, the observed variables are assumed to be statistically independent, conditional on the latent variables. Thus, the models assume that, given a specific level of a latent variable (e.g. depression), the indicators (e.g. 'feelings of guilt' and 'suicidal ideation') are uncorrelated. This feature, which is known as local independence, is consistent with a causal interpretation of the effects of the latent on the observed variables (Borsboom *et al.* 2003). In such a causal interpretation, variation in the latent variable is not merely associated with, but in fact causally responsible for, variation on the observed variables.

The distribution of observed variables is typically taken as a given in psychometric modelling, as it is dictated by the response format used in questionnaires or interviews. The structure and distribution of the latent variable, however, may feature as a research question, rather than a known. This is often the case in psychiatric nosology, because we do not have strong independent evidence to resolve the question of whether psychiatric disorders vary continuously or categorically in the population. In this case, one may apply the models in Table 1 in an attempt to determine the form of the latent structure. This can be done in two ways. First, by inspecting particular consequences of the model for specific statistical properties of (subsets of) items, such as the patterns of bivariate correlations expected to hold in the data (Waller & Meehl, 1998). Second, on the basis of global fit measures that allow one to compare whether a model with a categorical latent structure fits the observed data better than a model with a continuous latent structure (De Boeck *et al.* 2005; Lubke & Neale, 2006, 2008; Lubke & Miller, 2015). The former of these approaches is typically denoted by the name 'taxometrics', while the latter is not, and we will follow this terminology here; however, it should be noted that taxometrics rests on exactly the same psychometric model as general latent variable modelling, and in this sense the approaches are complementary (see also Schmitt *et al.* 2006; McGrath & Walters, 2012).

The logic underlying taxometric analysis is, at first sight, straightforward (but see Maraun *et al.* 2003;

Maraun & Slaney, 2005). If the underlying construct is continuous, then the covariance between any two indicators conditional on a given range of a proxy of the construct should be same regardless of the exact range. For example, if the variable underlying depression is a (single) continuous dimension, then the relation between the scores on, say, an insomnia item and a suicidal ideation item should be about the same in people with low, intermediate or high scores on a third item that plausibly acts as a proxy to the depression construct (e.g. a sad mood item). In contrast, if the underlying variable is a binary variable (comprising two classes, i.e. healthy *v*. depressed), then the covariance between any two indicators conditional on a given range of the proxy of the construct is expected to vary with the value of the proxy. Specifically, at low (high) values of the proxy, most individuals are healthy (depressed), i.e. they are in the healthy (depressed) class. Within this class the covariance among the items is expected to be zero (as per local independence). In contrast, at intermediate levels of the proxy, we will find a mixture of individuals that belong to either class. Here, the conditional covariance between the indicators is expected to be larger, because between-class differences contribute to the covariance. Taxometric analysis capitalizes on such implications of latent structure hypotheses.

To carry out a taxometric analysis, one arbitrarily chosen variable (e.g. mood) is denoted the 'index' variable, and is assumed to be a proxy for the underlying construct. Then, over a moving window of values on the index variable, the covariance between the other two variables (i.e. suicidal ideation and feelings of worthlessness) is plotted. If the underlying construct is categorical, the resulting covariance curve will be peaked at the point where the selected groups contain equally many individuals from each latent class: because groups with very low or very high scores on the index variable will be composed almost entirely of individuals from one latent class, in which we have local independence, the correlation between the plotted variables will be lower for very high or very low scores of the index variable. As a result, a peaked covariance curve suggests a categorical latent structure, while a flat curve suggests a continuous latent structure. This particular method is called MAXCOV (for MAXimum COVariance). In a similar vein, MAXEIG (MAXimum EIGenvalue) plots the eigenvalue of a matrix of item covariances instead of a simple bivariate covariance, allowing more variables to be used in the analysis. Several other methodologies have been constructed over the years, which use the same idea of identifying divergent predictions from different hypotheses concerning the structure of latent variables (Ruscio *et al.* 2006; McGrath & Walters, 2012).

Despite its popularity, it should be noted that the taxometric approach is not uncontroversial in psychometrics. The reason is that it has long rested purely on the visual inspection of a plotted function instead of on a formal hypothesis test (Haslam *et al.* 2012). Another reason is that one of its core assumptions (categorical latent structures will produce peaked covariance functions) is not necessarily true (Maraun *et al.* 2003; Maraun & Slaney, 2005); for example, violations of distributional assumptions concerning measurement error may lead to peaks and valleys in the covariance function even if the latent variable is continuous (Molenaar *et al.* 2010). Some of these concerns may be less of a problem in methods based on simulation (Ruscio *et al.* 2007), and McGrath & Walters (2012) suggest that, despite the above problems, taxometric procedures do perform reasonably well in systematic simulations. Nevertheless, the fact that taxometrics lacks a comprehensive mathematical foundation is a considerable weakness, because it implies that the validity of taxometric techniques must be judged on a case-by-case basis.

Complementary to taxometric analysis, one may use latent variable modelling as a framework in which to query the structure of psychiatric constructs. In such approaches, one can compare the fit of a model in which the latent variable is represented as being categorical with that of a model in which the latent variable is represented as a continuous dimension to decide which model is superior (Lubke & Miller, 2015). This counters at least some of the above concerns, as latent variable modelling approaches do rest on a firm mathematical basis (Lubke & Muthén, 2005). However, latent variable approaches are not without problems either. For instance, it is well known that many continuous variable models have statistically equivalent categorical or mixture counterparts: that is, a fitting model with a categorical latent variable does not imply that the construct itself is categorical, because a continuous model might fit the same data equally well (e.g. Molenaar & Von Eye, 1994; Halpin *et al.* 2011; see also Erosheva, 2005).

McGrath & Walters (2012) have systematically evaluated the performance of latent variable models and taxometric procedures, and propose a combination of modelling approaches, in which taxometric strategies are used to detect categorical structures, whereas latent class or profile models are used to select the optimal number of classes if the structure is determined to be categorical. A thoughtful combination of different methodologies indeed appears the most sensible currently available strategy for investigating the issue. However, it is remarkable that no systematic and principled methodological procedure appears to have emerged from the psychometric work on this issue.

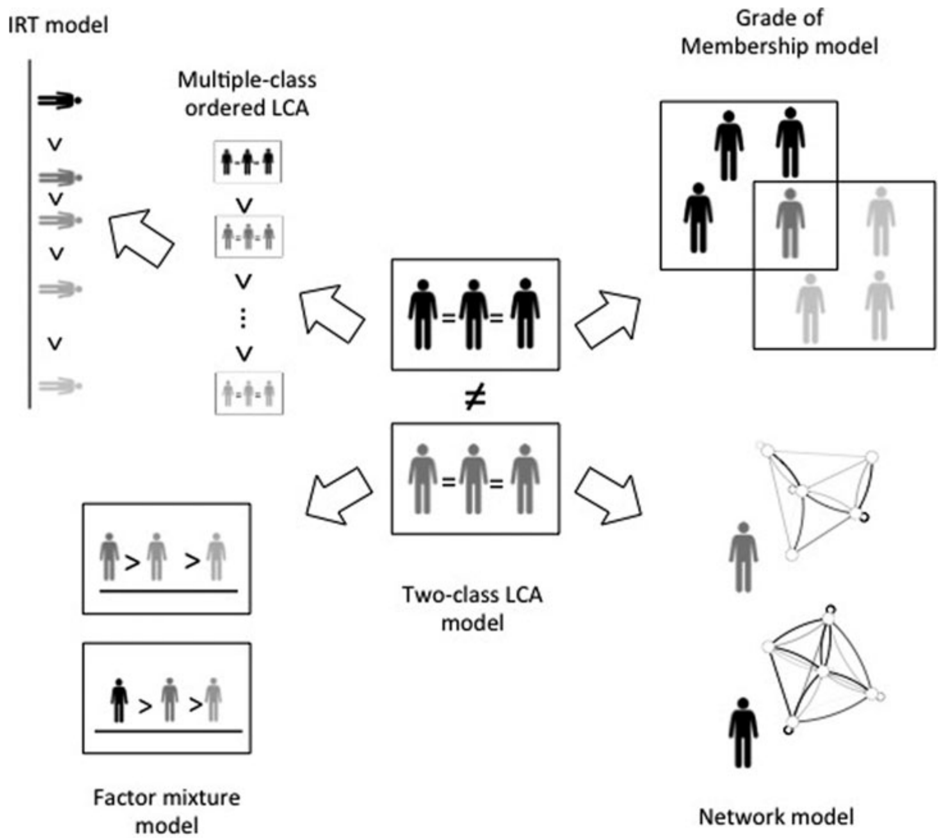

Fig. 2. Alternative psychometric models for representing kinds and continua. The left top panel show the standard progression from a two-class latent class analysis (LCA) to an item response theory (IRT) model, as in Fig. 1. The left bottom panel shows the inclusion of continuous variation within classes (factor mixture model). The right top shows a fuzzy set representation, in which individuals belong to multiple classes at the same time (grade-of-membership model). The right bottom panel shows network models, which can have categorical or continuous dynamics depending on the structure of the network (see Fig. 4).

One possible reason for this situation is that the hypotheses of kinds and continua do not exhaust the space of possibilities, so that evidence against one hypothesis is not necessary evidence for the other (as seems to be the implicit assumption in taxometrics). In the next sections, we discuss models that may indeed substantiate this idea, because they have both categorical and continuous aspects.

Alternative latent variable models

Recent developments in statistical modelling have produced various models that blur the kinds–continua distinction, in the sense that they accommodate both categorical and continuous latent structures at the same time. Fig. 2 illustrates how our basic model with two latent classes can be extended in two other directions than the standard order used in the foregoing [i.e. the transition from latent class analysis (LCA) to IRT as represented in Fig. 1]: by including continuous latent variables within classes (factor mixture models), and by making group membership itself a matter of degree [grade-of-membership (GoM) models]. In addition, network models allow the very same structure to be continuous and categorical, depending on the parameters of the network.

Factor mixture models

Finite mixture models partition the population into distinct latent classes, but allow for continuous variation within these classes (McLachlan & Peel, 2000). If that variation is itself measured through a number of indicator variables, then we obtain a factor mixture model (Yung, 1997; Dolan & Van der Maas, 1998; Arminger *et al.* 1999; Lubke & Muthén, 2005; Muthén, 2008). The factor mixture model can thus be understood as a latent class model in which each latent class is characterized by its own common factor model (see bottom left panel of Fig. 2). Importantly, conditioning on the latent classes no longer renders the observed variables independent, as their conditional

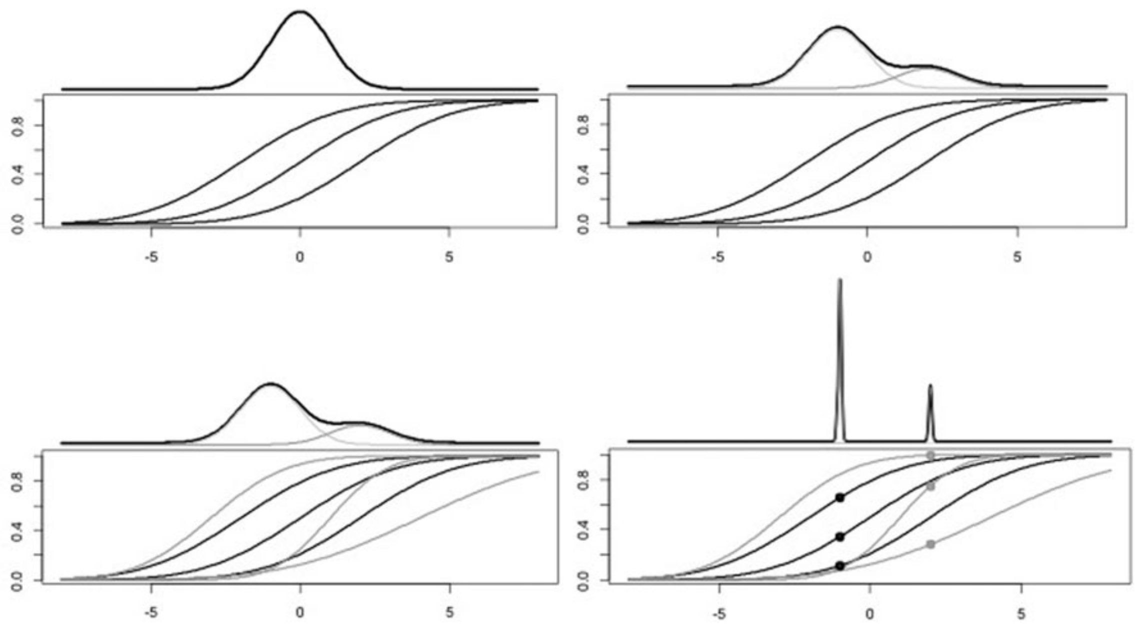

Fig. 3. From an item response theory model with a latent normal density (top left), via a two-component discrete factor mixture model subject to measurement invariance (top right) and the same model without measurement invariance (bottom left) to an unconstrained latent class model (bottom right).

distribution is characterized by the factor model. However, observed variables are assumed to be conditionally independent given both the latent class and the latent continuous factor. Alternatively, the factor mixture model can be understood as a multi-group common factor model in which group membership is unknown. The latent class variable then takes the place of an observed grouping variable. The factor mixture model may distinguish between healthy and affected individuals, but also allows for quantitative individual differences within the two classes. For instance, such a model may propose a categorical distinction between people with and without depression, but at the same time may allow for continuous variation in the degree to which depressive symptomatology is present within each of these classes.

Factor mixture models provide a useful framework for formalizing the distinction between categorical and continuous latent variables in terms of distributional assumptions and model constraints (Muthén, 2006; Masyn *et al.* 2010; Von Davier *et al.* 2012; Clark *et al.* 2013). For instance, consider Fig. 3, in which we display the latent distribution above the ICCs for three binary items, which specify how the expected value of the items (e.g. the expected value of the symptom 'suicidal ideation') varies with positions on the latent variable (e.g. major depression). The top-left figure represents an IRT model with the assumption of that the latent variable follows a normal distribution. One's position on the latent trait determines one's probability of endorsing the items through the ICCs in the figure directly below the normal density: e.g. the more depressed, the higher the probability of endorsing the item 'suicidal ideation'. This is a standard IRT model as discussed earlier.

Now we can gradually move from this dimensional model towards a categorical model by changing the model constraints. For example, the top-right figure represents the same model as the top-left figure, but now in two distinct groups which follow the same ICCs (so that measurement invariance holds; Mellenbergh, 1989; Meredith, 1993). The bottom-left figure represents this same two-group model, but here the groups are allowed to follow different ICCs. From this model, we can subsequently derive a latent class model by simply shrinking the variances in each group towards zero; here, the latent classes can be viewed as specific values on the continuous factor (under additional ordering restrictions, this model is equivalent to a model with ordered latent classes, as represented in the middle panel of Fig. 1; see also Croon, 1990; Vermunt, 2001; Masyn *et al.* 2010).

Thus, while the factor model and the latent class model are distinct models in Table 1, mixture modelling allows us to connect them by means of intermediate models and associated constraints. Importantly, the existence of these models implies that the choice between fully discrete and fully continuous latent structures does not exhaust the possibilities. As De Boeck *et al.* (2005) note, within this psychometric framework,

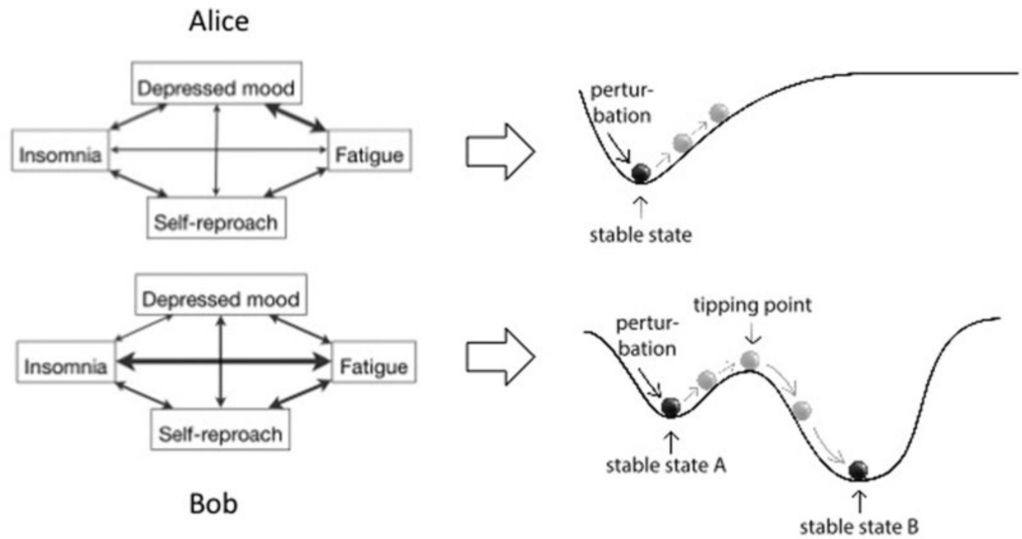

Fig. 4. Alice's network, which has a weak connectivity profile (left top) gives rise to a single basin of attraction (right top), corresponding to a healthy state. Continued stress (perturbation; force being exerted on the ball from the left) may cause prolonged changes in state (elevated ball), but upon removal of stress the network will return to the healthy state. Bob's more strongly connected network (bottom left) features alternative stable states (bottom right) and a tipping point. If external stress exceeds a given value, then the system collapses to the disordered state, which is itself stable. Thus, the system will maintain the disordered state, even if the stressor is removed.

the distinction between categories and continua is itself a matter of degree (see also Von Davier *et al.* 2012).

GoM models

In GoM models, one can also depart from a simple latent class model to integrate continuous features. Here, however, the continuous variation concerns group membership itself. Where the latent class model assumes that every person belongs to one and only one latent class, the GoM model allows individuals to be members of multiple classes at the same time (Erosheva, 2005). Thus, instead of proposing that an individual belongs either to, say, the class of typical or to the class of atypical depression (the focal hypothesis in LCA), the GoM model holds that an individual belongs to both classes at the same time, but to different degrees. The model thus is a psychometric instantiation of the logical concept of a 'fuzzy set': a set that has no clear boundaries, and to which different objects belong to different degrees (Verkuilen *et al.* 2011). In the GoM model, the degree of membership is expressed in terms of a set of probabilities that sum to unity (so, in a two-class model, if one's GoM for one class equals 60%, GoM to the other class must equal 40%). Erosheva (2005) and Asparouhov & Muthén (2008) provide treatments of different statistical representations of the GoM model[5]. The move from a latent class to a GoM model is illustrated in the top-right panel of Fig. 2.

The GoM model has been applied to schizophrenia (Manton *et al.* 1994; Jablensky, 2006, 2010), depression (Woodbury & Manton, 1989) and personality disorders (Nurnberg *et al.* 1999) and, in these applications, meaningful classes could indeed be defined. However the GoM model is not widely used in psychometric applications, probably due to the lack of readily accessible statistical software to apply the GoM model. However, this has recently changed with the appearance of Robitzsch's (2014) package for the statistical software environment R. The GoM model can also be fitted in M*plus* by representing it as a multilevel latent variable model (Asparouhov & Muthén, 2008). These software advances invite the further application of the model to clinical psychology and psychiatry.

Network models and dynamical systems

In traditional models discussed so far, theoretical constructs are assumed to be either categorical or continuous for all elements of the population. However, it is possible that the transition to and from a psychiatric disorder proceeds as a categorical sudden transition for some individuals, whereas it is a smooth process of change for others. Recently developed network models accommodate this possibility, and thus shed a new light on the question of whether disorders should be thought of as categories or as continua.

Psychometric latent variable models represent differences in the structure of psychiatric constructs as differences in the distributional form of a latent variable, which acts as a common cause of the indicators[6]. Thus, when one considers 'insomnia' and 'fatigue', two symptoms of depression, a latent variable model assumes that when insomnia and fatigue covary, this is the result of their common dependence on depression (insomnia←depression→fatigue). Recently developed network models of psychiatric disorders (Cramer *et al.* 2010, 2012*a*, *b*; Borsboom & Cramer, 2013), in contrast, assume that insomnia and fatigue covary because they are causally related: if one does not sleep, one will get tired eventually (insomnia → fatigue). Correlations between variables commonly seen as 'indicators' then arise from a network of causal effects among these variables themselves (i.e. they form so-called mechanistic property clusters; Kendler *et al.* 2011); as such, in networks, there are no latent variables that function as psychologically meaningful common causes, even though individual nodes and connections may of course stand under the influence of factors (e.g. genetic effects or life events) that are not directly observed.

Individual differences in network structure may lead to different patterns of symptom dynamics. For instance, insomnia may quickly result in fatigue in Bob's network (because of a strong connection between insomnia and fatigue), but less quickly in Alice's network (in which this connection is weaker; see Fig. 4). For a person with a weakly connected network, external stressors (like losing one's job) may lead to an increase in the number of symptoms that are activated but, importantly, when the external stressors are removed, the person will spontaneously and smoothly return to equilibrium. Strongly connected networks, however, can behave differently: they may show strongly non-linear behaviour with sudden jumps from one state to another (Thom, 1975; Zeeman, 1977; Van der Maas & Molenaar, 1992; Cramer, 2013). For example, for the case of depression this means that in a person with a strongly connected network, a given perturbation (e.g. an adverse life event) has the potential to trigger a full and persistent depressive episode, while that same perturbation may only yield a temporary elevation of depressed symptoms in a weakly connected network (see also Pe *et al.* 2015). The reason for this difference is that, in strongly connected networks, symptom activation may be increased through feedback loops; e.g. a person who suffers from lack of interest may avoid social contacts, which may lead to further diminished interest. Thus, in a strongly connected network, the spreading of activation through the symptom network can cause a depressive episode, which may then be maintained by the feedback loops in the network.

These differences in dynamics across different network structures are important to the kinds *v.* continua discussion, because they show that disorders may be discrete kinds for some people (e.g. people with strongly connected networks) and continuous structures for others (e.g. people with weakly connected networks). Thus, individual differences data may look like a continuous distribution even if, intra-individually, the transition from a healthy to a disordered state is discontinuous (see also Borsboom *et al.* 2003; Molenaar, 2004; Molenaar & Campbell, 2009; Adolf *et al.* 2014). This may explain why taxometric investigations of individual differences data often claim evidence for continuity, even in cases where the within-person phenomenology would seem to suggest a discontinuous shift from health to disorder, like post-traumatic stress (Forbes *et al.* 2005) and psychosis (Ahmed *et al.* 2012), and that evidence is often mixed, depending strongly on the specific sample that was investigated.

If present, discontinuous transitions have direct measureable consequences that may be exploited in further research, because transitions from a healthy state to a disordered state are typically preceded by early warning signals (EWS; Scheffer *et al.* 2009, 2012) that indicate that the system is close to a tipping point for a transition. An important EWS is critical slowing down, which means that a system near a transition will take longer to recover from random perturbations of its state. In the context of psychiatric disorders, this would mean that a person at the brink of developing an episode of depression will recover more slowly than a healthy person from a relatively minor daily stressors such as an unpleasant phone call with their mother-in-law. As a result, the system's state at a given time point becomes more predictable from previous time points. Studies on mood fluctuations indeed suggest that people who feature such increased predictability (emotional inertia; Kuppens *et al.* 2010) are at elevated risk for mood disorders, in line with the idea that mood shifts may be preceded by EWS (Van de Leemput *et al.* 2014) as the system becomes less resilient (see also Montpetit *et al.* 2010).

Thus, network models provide a fresh way of thinking about the problem of kinds *v.* continua, and suggest new avenues for research. Psychometrically, network models can be fitted as so-called Markov random fields (Kindermann & Snell, 1980; Epskamp *et al.* 2012). The R-packages IsingFit (Van Borkulo *et al.* 2014) and qgraph (Epskamp *et al.* 2012) implement data-analytic approaches that may be used to this end. Individual differences in network structures can be studied by using individual differences in time series, with a multilevel modelling framework that allows for inter-individual variation in intra-individual networks (Bringmann *et al.* 2013, 2015).

Discussion

The current evaluation of psychometric conceptualizations and models shows that the distinction between continua and kinds is considerably more subtle than has been presupposed in the literature. Several psychometric models explicitly integrate continuous and categorical aspects of the construct studied, and in network models the very same construct can show continuous or discontinuous behaviour, depending on the network structure. Novel psychometric models conceptualize the distinction between kinds and continua in interesting ways, and suggest new ways of thinking about the issue.

In particular, these approaches may serve to elucidate why, so far, few research lines have led to a univocal verdict on the structure of studied constructs. As Haslam *et al.* (2012) note, results vary over contexts, types of data, and substantive domains in ways that are not easily interpreted (see also Lubke & Miller, 2015). As a result, it is possible to find evidence both for and against taxa in the taxometric literature of almost any disorder. We have suggested here that one reason for this paucity of strong results may be that the models considered represent a limited and perhaps inappropriate subset of theoretical positions, which portray the kinds and continua hypotheses as mutually exclusive and exhaustive. A fair evaluation of the psychometric literature does not support this portrayal.

An important novel approach that may serve to increase our resolution concerning the structure of psychopathology involves the study of intra-individual transitions from and to disorder states (Boker *et al.* 2009; Molenaar *et al.* 2013). Ecological momentary assessment procedures provide a fruitful way of studying transitions (Hamaker *et al.* 2007; Van de Leemput *et al.* 2014). Wichers (2014) has described a theoretical program that connects the macro-level of mental disorders to the micro-dynamics of mood states as they fluctuate and influence each other over short time periods. The combination of this program with advanced psychometric models for time series is likely to shed light on the conditions that would lead to continuous *v*. categorical change.

A final extension that may further our understanding of categorical and continuous aspects of disorders concerns the domain of experimental manipulations. It is well known in the literature on phase transitions in other areas of psychology (Van der Maas *et al.* 1992) that the study of transitions requires a careful experimental setup, in which variables that drive transitions can be isolated and manipulated – ideally, the system is moved into and out of a number of distinct states within a single experiment (e.g. Dutilh *et al.* 2010). Of course, such manipulations are ethically questionable in clinical work; however, one could imagine clever combinations of quasi-experimental research (e.g. prospective studies involving adverse life events), experimental manipulations (e.g. treatment studies), and model-based simulations that might be used as proxies to a full experimentally controlled procedure. Incorporating such experimental features may considerably increase the resolution of psychometric attempts to uncover the structure of psychopathology.

In conclusion, both the models and the data used in current research on the structure of psychopathology are limited and often suboptimal. As a result, we should hesitate to draw strong conclusions from the current literature. Fortunately, extensions into models for within-person dynamics, experimental setups, and multivariate systems with continuous and categorical variables are now within reach. These developments are likely to shed more light on the structure of psychiatric constructs in the years to come.

Acknowledgements

D.B. is supported by ERC Consolidator Grant no. 647209. M.R. is supported by ERC Career Integration Grant no. 631145. A.O.J.C. is supported by Veni grant no. 016.155.083 from the Netherlands Organization for Scientific Research. M.S. is supported by a Spinoza Prize awarded by the Netherlands Organization for Scientific Research and ERC Advanced Grant no. 5120755-01.

Declaration of Interest

None.

Notes

[1] Not to be confused with test theory (Lord & Novick, 1968) or measure theory (Tao, 2011).

[2] Continuous structures may, but need not, be quantitative. In addition to continuity, quantitative attributes feature additive structure (Michell, 1997). For instance, distance is not merely continuous (for any two distances there exists a third in between), but also additively structured, because for any two distances a and b, if $a > b$ then we can find a third distance c such that $a = b + c$. Attributes that adhere to this requirement, and to several other axioms (Hölder, 1901) are not merely continuous but also quantitative. Such attributes can be represented on ratio or interval scales (Suppes & Zinnes, 1963; Krantz *et al.* 1971). It is often incorrectly assumed that continuous scales are necessarily quantitative, and occasionally the terms 'continuous' and 'quantitative' are even used interchangeably. However, a continuous scale need not have ratio of interval properties. For instance, the Dewey library classification system commonly used to archive books is

practically continuous, but carries only nominal information, as it serves merely to uniquely identify books (Markus & Borsboom, 2013).

[3] For continuous physical quantities (Michell, 1997; Cooper & Humphry, 2012), such as distance, the continuity hypothesis may be literally true. For instance, if you imagine all distances that would fit in the palm of your hand, no distance between zero and the size of your hand is missing from that set. In psychology, attributes cannot be continuous in this literal sense, because they are defined on the human population, which is finite. Thus, any psychological attribute will feature gaps in its actual structure. In this paper, we therefore take continuity to be an approximation to a (possible very large) number of ordered classes of actually instantiated values, or, in the case of intra-individual dynamical systems, to a set of possible values that a system would pass through if it were to move from one point on the scale to another.

[4] Note that a model with a single continuous latent variable does not merely hold that individuals vary in degree, but also that all psychometrically relevant information about these differences can be captured in a single number. Thus, the hypothesis formulated here is not merely that people differ from each other, but that that they differ in an extremely simple linearly ordered way. Formally, this hypothesis is known as the unidimensionality hypothesis.

[5] Technically, the GoM model is the mirror image of the factor mixture model: whereas the factor mixture model can be seen as an infinite mixture model (the factor model) nested in a finite mixture model (the class model), the GoM model can be seen as a finite mixture model (the class model) nested in an infinite mixture model (the factor model).

[6] In the language of the modern causality literature, local independence follows directly from the common cause interpretation of the latent variable, as the latter *d*-separates the observed variables from each other (Pearl, 2009).

References

Adolf J, Schuurman NK, Borkenau P, Borsboom D, Dolan CV (2014). Measurement invariance within and between individuals: a distinct problem in testing the equivalence of intra- and inter-individual model structures. *Frontiers in Quantitative Psychology and Measurement* **5**, 883.

Agresti A (2013). *Categorical Data Analysis*. Wiley: New York.

Ahmed AO, Buckley PF, Mabe PA (2012). Latent structure of psychotic experiences in the general population. *Acta Psychiatrica Scandinavia* **125**, 54–65.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. American Psychiatric Publishing: Arlington, VA.

Arminger G, Stein P, Wittenberg J (1999). Mixtures of conditional mean- and covariance-structure models. *Psychometrika* **64**, 475–494.

Asparouhov T, Muthén B (2008). Multilevel mixture models. In *Advances in Latent Variable Mixture Models* (ed. G. R. Hancock and K.M. Samuelsen), pp. 27–51. Information Age Publishing, Inc.: Charlotte, NC.

Bartholomew DJ (1987). *Latent Variable Models and Factor Analysis*. Griffin: London.

Boker SM, Molenaar PCM, Nesselroade JR (2009). Issues in intraindividual variability: individual differences in equilibria and dynamics over multiple time scales. *Psychology and Aging* **24**, 858–862.

Bollen KA (1989). *Structural Equations with Latent Variables*. Wiley: New York.

Borsboom D (2005). *Measuring the Mind: Conceptual Issues in Contemporary Psychometrics*. Cambridge University Press: Cambridge.

Borsboom D, Cramer AOJ (2013). Networks: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology* **9**, 91–121.

Borsboom D, Mellenbergh GJ, Van Heerden J (2003). The theoretical status of latent variables. *Psychological Review* **110**, 203–219.

Bringmann LF, Lemmens LHJM, Huibers MJH, Borsboom D, Tuerlinckx F (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Medicine* **45**, 747–757.

Bringmann LF, Vissers N, Wichers M, Geschwind N, Kuppens P, Peeters F, Borsboom D, Tuerlinckx F (2013). A network approach to psychopathology: new insights into clinical longitudinal data. *PLOS ONE* **8**, e60188.

Clark SL, Muthén B, Kaprio J, D'Onofrio BM, Viken R, Rose RJ (2013). Models and strategies for factor mixture analysis: an example concerning the structure underlying psychological disorders. *Structural Equation Modeling: A Multidisciplinary Journal* **20**, 681–703.

Cooper G, Humphry SM (2012). The ontological distinction between units and entities. *Synthese* **187**, 393–401.

Cramer AOJ (2013). The glue of (ab)normal mental life: networks of interacting thoughts, feelings and behaviors. Ph.D. Thesis (http://dare.uva.nl/record/452479). Accessed September 2015.

Cramer AOJ, Borsboom D, Aggen SH, Kendler KS (2012*a*). The pathoplasticity of dysphoric episodes: differential impact of stressful life events on the patterns of depressive symptom inter-correlations. *Psychological Medicine* **42**, 957–965.

Cramer AOJ, van der Sluis S, Noordhof A, Wichers M, Geschwind N, Aggen SH, Kendler KS, Borsboom D (2012*b*). 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 AOJ, Waldorp LJ, van der Maas HLJ, Borsboom D (2010). Comorbidity: a network perspective. *Behavioral and Brain Sciences* **33**, 137–193.

Cronbach LJ, Meehl PE (1955). Construct validity in psychological tests. *Psychological Bulletin* **52**, 281–302.

Croon MA (1990). Latent class analysis with ordered latent classes. *British Journal of Mathematical and Statistical Psychology* **43**, 171–192.

De Boeck P, Wilson M, Acton GS (2005). A conceptual and psychometric framework for distinguishing categories and dimensions. *Psychological Review* **112**, 129–158.

Dolan CV, Van der Maas HLJ (1998). Fitting multivariate normal finite mixtures subject to structural equation modeling. *Psychometrika* **63**, 227–253.

Dutilh G, Wagenmakers EJ, Visser I, Van der Maas HLJ (2010). A phase transition model for the speed–accuracy trade-off in response time experiments. *Cognitive Science* **34**, 211–250.

Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software* **48**, 1–18.

Erosheva EA (2005). Comparing latent structures of the grade of membership, Rasch, and latent class models. *Psychometrika* **70**, 619–628.

Forbes D, Haslam N, Williams BJ, Creamer M (2005). Testing the latent structure of posttraumatic stress disorder: a taxometric study of combat veterans. *Journal of Traumatic Stress* **18**, 647–656.

Halpin PF, Dolan CV, Grasman RPPP, De Boeck P (2011). On the relation between the linear factor model and the latent profile model. *Psychometrika* **76**, 564–583.

Hamaker EL, Nesselroade JR, Molenaar CM (2007). The integrated trait–state model. *Journal of Research in Personality* **41**, 295–315.

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

Hölder O (1901). Die Axiome der Quantität und die Lehre vom Mass (The axioms of quantity and the doctrine of weight). *Ber. Verh. Kgl. Sächsis. Ges. Wiss. Leipzig, Math.-Phys. Classe* **53**, 1–64.

Hyland ME (2011). *The Origins of Health and Disease*. Cambridge University Press: Cambridge, UK.

Jablensky A (2006). Subtyping schizophrenia: implications for genetic research. *Molecular Psychiatry* **11**, 815–836.

Jablensky A (2010). The diagnostic concept of schizophrenia: its history, evolution, and future prospects. *Dialogues in Clinical Neuroscience* **12**, 271–287.

Kendler KS, Zachar P, Craver C (2011). What kinds of things are psychiatric disorders? *Psychological Medicine* **41**, 1143–1150.

Kindermann R, Snell JL (1980). *Markov Random Fields and their Applications*. American Mathematical Society: Providence, RI.

Kraepelin E, Dierendorf AR (1915). *Clinical Psychiatry. A Textbook for Students and Physicians*. The MacMillan Company: New York.

Krantz DH, Luce RD, Suppes P, Tversky A (1971). *Foundations of Measurement*, vol. **I**. Academic Press: New York.

Kuppens P, Allen NB, Sheeber LB (2010). Emotional inertia and psychological maladjustment. *Psychological Science* **21**, 984–991.

Lazarsfeld PF, Henry NW (1968). *Latent Structure Analysis*. Houghton-Mifflin: Boston, MA.

Lord FM, Novick MR (1968). *Statistical Theories of Mental Test Scores*. Addison-Welsley: Reading, MA.

Lubke GH, Miller PJ (2015). Does nature have joints worth carving? A discussion of taxometrics, model-based clustering and latent variable mixture modeling. *Psychological Medicine* **45**, 705–715.

Lubke GH, Muthén B (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods* **10**, 21–39.

Lubke GH, Neale M (2006). Distinguishing between latent classes and continuous factors: resolution by maximum likelihood. *Multivariate Behavioral Research* **41**, 499–532.

Lubke GH, Neale M (2008). Distinguishing between latent classes and continuous factors with categorical outcomes: class invariance of parameters of factor mixture models. *Multivariate Behavioral Research* **43**, 592–620.

MacCallum RC, Zhang S, Preacher KJ, Rucker DD (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods* **7**, 19–40.

Manton KG, Korten A, Woodbury MA, Anker M, Jablensky A (1994). Symptom profiles of psychiatric disorders based on graded disease classes: an illustration using data from the WHO International Pilot Study of Schizophrenia. *Psychological Medicine* **24**, 133–144.

Maraun MD, Slaney K (2005). An analysis of Meehl's MAXCOV-HITMAX procedure for continuous indicators. *Multivariate Behavioral Research* **40**, 489–518.

Maraun MD, Slaney K, Goddyn L (2003). An analysis of Meehl's MAXCOV-HITMAX procedure for dichotomous indicators. *Multivariate Behavioral Research* **38**, 81–112.

Markus KA, Borsboom D (2012). The cat came back: evaluating arguments against psychological measurement. *Theory and Psychology* **22**, 452–466.

Markus KA, Borsboom D (2013). *Frontiers of Test Validity Theory: Measurement, Causation, and Meaning*. Routledge: New York.

Masyn K, Henderson C, Greenbaum P (2010). Exploring the latent structures of psychological constructs in social development using the dimensional–categorical spectrum. *Social Development* **19**, 470–493.

McCutcheon AL (1987). *Latent Class Analysis*. *Quantitative Applications in the Social Sciences Series* no. 64. Sage: Thousand Oaks, CA.

McGrath RE, Walters GD (2012). Taxometric analysis as a general strategy for distinguishing categorical from dimensional latent structure. *Psychological Methods* **17**, 284–293.

McLachlan G, Peel D (2000). *Finite Mixture Models*. Wiley: New York.

Meehl PE (1992). Factors and taxa, traits and types, difference of degree and differences in kind. *Journal of Personality* **60**, 117–174.

Meehl PE (1995). Bootstraps taxometrics: solving the classification problem in psychopathology. *American Psychologist* **50**, 266–275.

Mellenbergh GJ (1989). Item bias and item response theory. *International Journal of Educational Research* **13**, 127–143.

Mellenbergh GJ (1994). Generalized linear item response theory. *Psychological Bulletin* **115**, 300–307.

Meredith W (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika* **58**, 525–543.

Michell J (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology* **88**, 355–383.

Michell J (1999). *Measurement in Psychology: A Critical History of a Methodological Concept*. Cambridge University Press: Cambridge, UK.

Molenaar D, Dolan CV, Verhelst ND (2010). Testing and modeling non-normality within the one factor model. *British Journal of Mathematical and Statistical Psychology* **63**, 293–317.

Molenaar PCM (2004). A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives* **2**, 201–218.

Molenaar PCM, Campbell CG (2009). The new person-specific paradigm in psychology. *Current Directions in Psychology* **18**, 112–117.

Molenaar PCM, Lerner RM, Newell KM (2013). *Handbook of Developmental Systems*. Guilford: New York.

Molenaar PCM, Von Eye A (1994). On the arbitrary nature of latent variables. In *Latent Variables Analysis* (ed. A. Von Eye and C.C. Clogg), pp. 226–242. Sage Publications: Thousand Oaks, CA.

Montpetit MA, Bergeman CS, Deboeck PR, Tiberio SS, Boker SM (2010). Resilience-as-process: negative affect, stress, and coupled dynamical systems. *Psychology and Aging* **25**, 631–640.

Muthén B (2006). Should substance use disorders be considered as categorical or dimensional? *Addiction* **101** (Suppl. 1), 6–16.

Muthén B (2008). Latent variable hybrids: overview of old and new models. In *Advances in Latent Variable Mixture Models* (ed. G. R. Hancock and K. M. Samuelsen), pp. 1–24. Information Age: Charlotte, NC.

Nurnberg HG, Woodbury MA, Bogenschutz MP (1999). A mathematical typology analysis of DSM-III-R personality disorder. *Comprehensive Psychiatry* **40**, 61–71.

Pe ML, Kircanski K, Thompson RJ, Bringmann LF, Tuerlinckx F, Mestdagh M, Mata J, Jaeggi SM, Buschkuehl M, Jonides J, Kuppens P, Gotlib IH (2015). Emotion-network density in major depressive disorder. *Clinical Psychological Science* **3**, 292–300.

Pearl J (2009). *Causality: Models, Reasoning, and Inference*, 2nd edn. Cambridge University Press: Cambridge, UK.

Reise SP, Waller NG (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology* **5**, 27–48.

Robitzsch A (2014). sirt: Supplementary Item Response Theory Models. R package version 0.47–36 (http://cran.r-project.org/web/packages/sirt/). Accessed September 2015.

Ruscio J, Haslam N, Ruscio AM (2006). *Introduction to the Taxometric Method: A Practical Guide*. Lawrence Erlbaum Associates: Mahwah, NJ.

Ruscio J, Ruscio AM, Meron M (2007). Applying the bootstrap to taxometric analysis: generating empirical sampling distributions to help interpret results. *Multivariate Behavioral Research* **42**, 349–386.

Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V, Held H, van Nes EH, Rietkerk M, Sugihara G (2009). Early-warning signals for critical transitions. *Nature* **461**, 53–59.

Scheffer M, Carpenter SR, Lenton TM, Bascompte J, Brock W, Dakos V, van de Koppel J, van de Leemput IA, Levin SA, van Nes EH, Pascual M, Vandermeer J (2012). Anticipating critical transitions. *Science* **338**, 344–348.

Schmitt JE, Aggen SH, Mehta PD, Kubarych TS, Neale MC (2006). Semi-nonparametric methods for detecting latent non-normality: a fusion of latent trait and ordered latent class modeling. *Multivariate Behavioral Research* **41**, 427–443.

Stevens SS (1946). On the theory of scales of measurement. *Science* **103**, 667–680.

Suppes P, Zinnes JL (1963). Basic measurement theory. In *Handbook of Mathematical Psychology* (ed. R.D. Luce, R. Bush and E. Galanter), pp. 3–76. Wiley: New York.

Tao T (2011). *An Introduction to Measure Theory*. American Mathematical Society: Providence, RI.

Thom R (1975). *Structural Stability and Morphogenesis*. Benjamin Press: Reading, MA.

Trendler G (2009). Measurement theory, psychology, and the revolution that cannot happen. *Theory and Psychology* **19**, 579–599.

Van Borkulo CD, Borsboom D, Epskamp S, Blanken TF, Boschloo L, Schoevers RA, Waldorp LJ (2014). A new method for constructing networks from binary data. *Scientific Reports* **4**, 5918.

Van de Leemput IA, Wichers M, Cramer AOJ, Borsboom D, Tuerlinckx F, Kuppens P, Van Nes EH, Viechtbauer W, Giltay EJ, Aggen SH, Derom C, Jacobs N, Kendler KS, Van der Maas HLJ, Neale MC, Peeters F, Thiery E, Zachar P, Scheffer M (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences USA* **111**, 87–92.

Van der Maas HLJ, Molenaar PCM (1992). Stagewise cognitive development: an application of catastrophe theory. *Psychological Review* **99**, 395–417.

Van der Sluis S, Posthuma D, Nivard MG, Verhage M, Dolan CV (2013). Power in GWAS: lifting the curse of the clinical cut-off. *Molecular Psychiatry* **18**, 2–3.

Verkuilen J, Kievit RA, Zand Scholten A (2011). Representing concepts by fuzzy sets. In *Concepts and Fuzzy Logic* (ed. R. Belohavek and G.J. Klir), pp. 149–176. MIT Press: Cambridge, MA.

Vermunt JK (2001). The use of restricted latent class models for defining and testing nonparametric and parametric item response theory models. *Applied Psychological Measurement* **25**, 283–294.

Von Davier M, Naemi B, Roberts RD (2012). Factorial *versus* typological models: a comparison of methods for personality data. *Measurement: Interdisciplinary Research and Perspectives* **10**, 185–208.

Waller NG, Meehl PE (1998). *Multivariate Taxometric Procedures: Distinguishing Types from Continua*. Sage: Thousand Oaks, CA.

Wichers M (2014). The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine* **44**, 1349–1360.

Woodbury MA, Manton KG (1989). Grade of membership analysis of depression-related psychiatric disorders. *Sociological Methods and Research* **18**, 126–163.

World Health Organization (1992). *International Statistical Classification of Diseases, Injuries, and Causes of Death. Sixth Revision of the International List of Diseases and Causes of Death*. World Health Organization: Geneva.

Yung YF (1997). Finite mixtures in confirmatory factor-analysis models. *Psychometrika* **62**, 297–330.

Zeeman EC (1977). *Catastrophe Theory: Selected Papers*. Addison-Wesley: Reading, MA.