The Challenges Raised by Comorbidity in Psychiatric Research: The Case of Autism

Valentina Petrolini & Agustín Vicente

Lindy Lab, University of the Basque Country (UPV/EHU) & Ikerbasque

Forthcoming Philosophical Psychology

Abstract

Despite several criticisms surrounding the DSM classification in psychiatry, a significant bulk of research on mental conditions still operates according to two core assumptions: a) *homogeneity*, that is the idea that mental conditions are sufficiently homogeneous to justify generalization; b) *additive comorbidity*, that is the idea that the coexistence of multiple conditions in the same individual can be interpreted as additive. In this paper we take autism research as a case study to show that, despite a plethora of criticism, psychiatric research often continues to operate in accordance with this model. Then we argue that such a model runs into problems once facts about *comorbidity* are taken into account. Finally, we offer some suggestions on how to tackle the challenge raised by comorbidity and its impact on heterogeneity. To do so, we explore *transdiagnostic stratification* accounts and *network models* to show that combining these approaches can move us in the right direction.

Keywords

DSM-5; comorbidity; heterogeneity; complexity; autism; stratification; transdiagnostic

Introduction

The *Diagnostic and Statistical Manual of Mental Disorders* (DSM henceforth), an instrument widely used to assess and diagnose mental conditions worldwide, has been criticized on various

fronts by clinicians and researchers alike. Most of these criticisms, which we describe in more detail in the next section, are centered around the categorical, polythetic, and descriptive nature of the manual. In this paper we develop a new line of criticism, using autism as a case study. Our critique takes a conservative approach towards the DSM. The argument starts out by granting that the DSM tracks conditions pertaining to individuals, *and* that such conditions rarely come alone. From this conservative viewpoint, we establish the following: (i) facts about comorbidity show that exhibiting more than one DSM condition is the norm rather than the exception; (ii) symptoms pertaining to different DSM conditions interact in non-additive, complex, and often unpredictable ways. Accepting (i) and (ii) has at least two unfortunate consequences for the DSM and for the research conducted following its guidelines. First, our argument shows that the *heterogeneity* of mental conditions - already acknowledged by most authors - gets drastically exacerbated once comorbidity is taken into account. Second, the DSM will rarely be describing individuals in a way that is meaningful for clinical research.

§1. The DSM and its discontents

The nature of mental conditions (or disorders) is hotly debated in psychiatry and philosophy of psychiatry. Over the past few decades most criticisms have clustered around symptom-based diagnostic methods, epitomized by the various editions of the DSM. Although the explicit focus of the DSM has always been on practical goals - e.g., facilitating communication among clinicians, improving diagnostic reliability, and structuring pharmaceutical research or funding proposals (Tabb 2020; Cooper 2014; First 2012) - its pervasive use in clinical practice has prompted a number of theoretical reflections. Specifically, three features of the DSM have been the object of considerable debate: its implicit commitment to a *categorical* view of mental

conditions; its *polythetic* structure; and its *descriptive* approach (see Fellowes 2021; Lilienfeld & Treadway 2016; Cooper 2005 & 2020 just for a few examples of such discussions).[1]

With respect to *categoricity*, the DSM never explicitly endorsed the idea that mental conditions should be seen as discrete entities with sharp boundaries between each other and healthy states. However, clinicians and researchers in psychology and psychiatry have often defended some version of the categorical view, which regards mental conditions as categories with distinct etiologies separating them from normal states and from one another. In this respect, it is illustrative to look at the exchanges surrounding the development and publication of the DSM-5 (APA 2013), which started out with the prospect of prominently including dimensional measures (Kupfer et al., 2002; Regier et al., 2010) but ended up largely retaining traditional categories (Blashfield et al., 2014; Widiger & Crego 2015). Philosophers of psychiatry have also defended various versions of the medical model by adopting less radically categorical positions, such as accounts centered on prototypes and exemplars (Murphy 2006) and characterizations of mental conditions as homeostatic or mechanistic property clusters underwritten by a set of more or less stable biological mechanisms (Tsou 2016; Samuels 2009). The categorical view of mental conditions has attracted a plethora of criticism from philosophers and clinicians (Zachar 2000; Haslam 2014). Some argue that DSM categories have been unduly reified (Hyman 2010; Kendell & Jablensky 2003), while others claim that they fail to reflect relevant neurobiological or behavioral systems (Cuthbert & Insel 2013).

[1] There are other unfortunate consequences of the DSM structure which we do not directly discuss here. One of them is the large number of Not Otherwise Specified (NOS) diagnoses, which are bound to exclude people who exhibit anomalous profiles and/or people who score right below the relevant diagnostic threshold (Lilienfeld & Treadway 2016).

With respect to the DSM's *polythetic* structure, the main point of contention concerns the fact that most diagnostic criteria allow individuals to cross the clinical threshold(s) in different ways. In fact, the majority of diagnoses included in the DSM-5 (i.e., 58,3%) allow people to be classified as having the same condition without sharing any symptoms (Olbert et al., 2014). Examples include the checklist for Major Depression, which requires the individual to meet at least six diagnostic criteria out of nine, or Autism Spectrum Disorder which notoriously encompasses a wide range of individuals with heterogeneous profiles (more on this in §3). In some extreme cases, such as Borderline Personality Disorder, there might be more than one hundred ways to meet the relevant diagnostic criteria (Lenzenweger 2010). An important consequence of the DSM's polythetic structure is that it allows for significant *heterogeneity* so that people who share a psychiatric diagnosis often have very little in common (Allsopp et al., 2019).[2]

With respect to the *descriptive* approach championed by the DSM, psychiatrists tend to approach diagnosis pragmatically, by identifying a condition based on what causes the greater amount of suffering to the patient, or on the symptoms that better respond to treatment (Maj 2011). This is often seen as a consequence of a great causal complexity in psychiatry, where the presence of distinct disease entities is difficult to establish due to insufficient knowledge about etiology, pathophysiology, and underlying mechanistic explanations (Pinkus, Tew & First 2004; Maung 2016). Within somatic medicine, comorbidity is characterized as the simultaneous occurrence of two or more diseases, with distinct etiology or pathogenesis, in the same individual (Vella, Aragona & Alliani 2000; Feinstein 1970). Clinicians usually employ this notion to distinguish

[2] On some readings - see for instance Fellowes 2021 - polythetic diagnoses are regarded as useful *exactly* because they make room for significant heterogeneity. We do not engage with this point directly here, as we are more interested in criticizing the standard approach as a whole, but it is worth mentioning that heterogeneity has at times been described as a positive feature of the current model.

between a patient's "primary condition" - i.e. the one reflected by the core diagnosis (e.g., type 2 diabetes) - and concomitant or concurrent ones (e.g., hypertension, cardiovascular disease). Given that compelling etiological explanations in psychiatry are few and far between, the very distinction between primary conditions and comorbidities often ends up being subject to a certain degree of arbitrariness instead of being grounded on reliable criteria – e.g., key neurological disturbances, causal or mechanistic priority, etc. (Vella, Aragona & Alliani 2000). As a result, in most cases it is difficult to understand whether concomitant diagnoses reflect the presence of distinct clinical entities or refer to multiple manifestations of a single clinical entity (Maj 2005). Staggering rates of *psychiatric comorbidity* are one of the unfortunate consequences of such an approach (Aragona 2009a; van Loo & Romeijn 2015), combined with the categorical tendency towards "splitting" diagnostic entities into several narrowly-defined disorders (Aragona 2009b; First 2005).

In response to all these issues, several modifications of the DSM have been implemented and alternatives have been proposed. Some criticisms of the categorical model successfully found their way into more recent editions of the manual itself (APA 2013), such as the unification of three previous categorical diagnoses of autism into a single spectrum (Kapp & Ne'eman 2020), and the "Alternative DSM-5 Model for Personality Disorders" published in Section III of the DSM-5 (see Blashfield et al., 2014; Widiger & Crego 2015 on the controversies surrounding this decision). Criticisms about categoricity also brought about some promising alternatives in terms of research frameworks. One prominent example is the *Research Domain Criteria* project (RDoC) promoted by the National Institute for Mental Health (NIMH) in the United States, which aims at characterizing mental conditions along multiple domains and dimensions of functioning by drawing on genetic as well as behavioral evidence (Insel et al., 2010; Cuthbert

2014). Other research programs geared towards a dimensional and quantitative approach to psychopathology - such as the *Roadmap for Mental Health Research* (ROAMER, Schumann et al., 2014) and the *Hierarchical Taxonomy of Psychopathology* (HiTOP, Kotov et al., 2017) - broaden the focus beyond intra-individual factors and include domains related to public health, infrastructures, and socio-environmental components. Recently, network models of psychopathology (Borsboom & Cramer 2013; Borsboom et al., 2019) have also gained traction as a promising alternative to the dominant approach, in particular with respect to the view of symptoms as signs of underlying pathological disruptions. By contrast, network models propose to reconceptualize mental disorders as complex networks of interacting symptoms (Borsboom 2017; Fried et al., 2017).

Criticisms of the DSM model also come from radically alternative approaches that call into question the very idea that mental conditions could be understood in isolation as something that pertains exclusively to individuals. As social models of disability have argued most prominently, individuals present differences or divergences that combine with environmental, material, and social factors or constraints thereby generating specific challenges and - in some cases - behavioral conditions that are classified as medical disabilities (Chapman 2020 & 2019; Bervoets & Hens 2020). On these views, considerations on whether a given collection of traits or behavior counts as disordered are importantly dependent on value-laden criteria concerning what is socially acceptable, desirable, deviant, and so on (Cooper 2020; Verhoeff 2013). As a consequence, mental conditions are not characterized as properties that can be properly attributed to individuals alone, but rather as the result of a complex mismatch between individual characteristics, features of the environment, and social, material, or cultural structures (Milton 2016).

We mention these alternatives here to acknowledge that the assumptions underlying the DSM have been heavily criticized by researchers both in psychiatry and philosophy, leading to a range of different approaches to classification and research. We will discuss some of these approaches in more detail in §3. However, for now it is important to stress that - despite these crucial lines of criticism - a significant bulk of research in psychiatry keeps proceeding in accordance with the model outlined above. In the next section we show that this is the case by exploring two core assumptions - *homogeneity* and *additivity* - which are still widespread in psychiatry (§2). Specifically, we do so by discussing the case of *autism* research,[3] where these two assumptions are overwhelmingly employed as the default starting point to set up case-control studies and to investigate comorbidities. In §3 we take a closer look at comorbidity, and in particular at how symptoms that belong to different conditions - at least in principle - interact with one another. Then we develop our argument from comorbidity which adds up to recent criticisms of DSM classifications. Basically, even if we were to assume that DSM conditions track actual conditions that individuals exhibit, given the high rates of comorbidities and the way in which such conditions interact, the actual variety of symptoms exhibited by the majority of individuals will fail to correspond to any DSM category (or to any combination of them). Hence, since DSM categories at best only apply to a minority of individuals, we suggest a move towards a more transdiagnostic approach to mental and developmental conditions.

[3] In order to avoid pathologization, in this paper we use the labels 'autistic spectrum conditions' (ASC) and 'autism' interchangeably. Moreover, following the results reported in recent qualitative analyses of linguistic policies surrounding autism (Kenny et al. 2016; Botha et al. 2020), we use identity-first language in our discussion (i.e., "autistic" instead of "person with autism"). However, we acknowledge that 'autism spectrum disorder' (ASD) is still widely used as a diagnostic label (APA 2013), and that ASC is not immune to criticism within the autistic community (Bervoets & Hens 2020). We therefore use ASD only when directly referring to existing studies where such a label is employed.

§2. The standard approach: homogeneity and additivity

Despite the theoretical and nosological complexity discussed above, research on mental conditions still appears to rest on two core assumptions: a) *Homogeneity*: members of the same class - that is, individuals sharing the same diagnosis - are sufficiently homogeneous to support generalizations; b) *Additivity*: interactions between members of different classes - e.g., individuals with multiple diagnoses - are usually interpreted as additive.

These two assumptions can be independently questioned. The former concerns the idea that mental conditions may be identified as constitutive elements that are sufficiently homogeneous to ground relevant generalizations concerning diagnostic profiles, treatment, etc. In other words, by studying groups of individuals who exhibit a given condition, researchers aim to draw general conclusions that can be applied to most individuals with the same diagnosis (or at least to a significant portion of them). The latter assumption builds on the former and concerns what happens when different constitutive elements interact with one another. In cases of comorbidity, the interaction between such conditions tends to be interpreted as additive. The notion of additivity comes from debates concerning emergentism in complex systems (see Kauffman, 1995; Kim, 2006, among others). Defenders of emergentism hold that whole systems have properties that cannot be fully explained or predicted on the basis of the known causal powers of their basal conditions. By contrast, deniers of emergency maintain that the whole system's behavior can always be predicted (at least in principle) from the causal powers of its constituents. For instance, we can predict the causal powers of a 9-kg object made out of three objects weighing 3 kgs each. Although the 9-kg object will be able to do things none of the constituting objects do, we can predict its powers just by knowing about 3-kg objects and how their causal powers add up. Additivity thus consists in the assumption that causal powers of whole systems

can be always explained this way. In the case of psychiatric research, the assumption translates into the idea that we can get to know about multiple conditions exhibited by individuals *if* we know about each of them separately. In other words, additivity predicts that individuals with comorbidities (A+B) will exhibit symptoms of condition A *plus* symptoms of condition B. Researchers often aim to understand cases of comorbidity by projecting from cases where individuals exhibit one condition or another, assuming that conditions have robustly similar effects (i.e., regardless of the context in which they appear).

In this paper we argue that the assumptions of homogeneity and additivity are problematic. Yet, unlike the more radical criticism put forward by, e.g., social models of disability, we want to show that these assumptions should be called into question *even if we accept a more conservative starting point* about individuals and mental conditions. That is, even if we do not reject from the outset the idea that the DSM classifications are tracking conditions which are inherent to individuals - as more radical proposals do - heterogeneity and comorbidity raise important challenges to the current approach. In the next subsection we discuss some prominent examples from autism research to show that psychiatric research *de facto* continues to operate in accord with the assumptions of homogeneity and additivity.

§2.1. The case-control paradigm

We take the dominant approach in psychiatry to be an implicit endorsement of the model described above, one that reflects how a significant bulk of research is still conducted. Here we focus on the standard practice of *case-control studies* in autism research to exemplify this point. We are not arguing that the approach outlined above describes psychiatric research as a whole. In fact, we are aware that there are several exceptions to this paradigm. Yet, for our main claim to

pass muster it is sufficient to show that assumptions about homogeneity and additive comorbidity are still significantly widespread (or implicitly endorsed) in a number of studies on mental conditions.

Classic case-control studies still represent the golden standard in psychiatric research (Lombardo et al., 2019). They are usually built to set up a comparison between two groups, one comprising subjects who received a diagnosis for a mental condition (e.g., ASD) and the other including non-clinical subjects - i.e., people who never received a psychiatric diagnosis. In this respect, the case-control paradigm exemplifies a 'one-size-fits-all' approach where all cases are treated identically due to the same diagnostic label (Lombardo et al., 2019, p. 1439). Generally speaking, these studies are built around the idea that by comparing a group with a condition to a group without such a condition it is possible to reach a more general conclusion concerning the relevant condition. This in turn rests on the assumption that members of the relevant categories are sufficiently *homogeneous*.

The literature on autism offers several examples of the case-control methodology, where autistic and neurotypical participants are compared along several dimensions. Some of the most prominent subfields include studies on implicatures (Pijnaker et al., 2009), analogical reasoning (Morsanyi et al. 2019), theory of mind (Baron-Cohen 1989; Brewer, Young & Barnett 2017), and executive functions (Yang et al., 2009; Corbett et al., 2009). More often than not, the results reported in these studies tend to be mixed and at odds with one another. For instance, some researchers report that the autistic profile does not exhibit specific issues with certain types of metaphors and other similar figures of speech (Giora et al., 2012; Kasirer & Mashal, 2016), as well as with some implicatures, or with irony when accompanied by cues like prosody (Chevallier et al., 2011). By contrast, evidence from other studies shows that, in general, non-

literal meaning is an issue in autism (see Chahboun et al., 2017; Morsanyi et al., 2020, for metaphors, and Vulchanova et al., 2015; Walenski & Love, 2017, for idioms).

However, meta-reviews reporting such mixed evidence tend to reach conclusions concerning whether or not a given feature is typical in the relevant population. For instance, in a recent meta-review on analogical reasoning in autism, Morsanyi et al. (2019) first acknowledge that the empirical evidence is mixed, ranging from reports of impairment in autism, to no group differences, or even outstanding performance for autistic participants. The researchers then move on to declare that, by extracting data from multiple studies and combining them in a meta-analysis, their aim is to obtain a clearer picture regarding the presence or absence of group differences in analogical reasoning (p. 67). After excluding studies who did not match typically developing and autistic groups on chronological age and IQ, Morsanyi and colleagues conclude that analogical cognition is not affected in autistic individuals (as such). Notably, the usual motivation for matching groups on a factor such as IQ is to exclude confounding factors so that researchers are in a position to assess the actual capacities related to the condition they want to study.

Yet, there may be a lingering suspicion that even these more classical case-control studies implicitly explain some features of a given group by reference to some other co-occurring condition. For instance, by excluding those studies that did not match groups on IQ, Morsanyi and colleagues suggest that the poorer performance of autistic individuals on analogical reasoning tasks might be due to a co-occurring intellectual disability. However, we should be wary of such an inference. The fact that results *X* appear in groups exhibiting condition C (i.e, autism) and some other comorbid condition COM (i.e., intellectual disability) does not imply that the co-occurring condition should be seen as responsible for the relevant behavior. As we

suggest below (§3.2), it is rather likely that *X* relate to the interaction between the two conditions (i.e., C and COM).

When the effect of comorbid factors is explicitly taken into account, the design of case-control studies usually includes the comparison of three (or more) groups. A group of people diagnosed with condition C without any specific comorbidity (C-COM) is compared with a group of people diagnosed with C plus said comorbid condition (C+COM), and with a group of people who exhibit the comorbid condition only (COM). If the behavior of the C+COM group is more similar to the COM group than to the C-COM group, the trait in question is attributed to COM as opposed to the primary condition under consideration.[4] Take ASD and Attention Deficit Hyperactivity Disorder (ADHD): these two conditions exhibit high rates of comorbidity - i.e., up to 40-50% in 3-12 year old children - and are directly compared in several case-control studies. For instance, as reported by Matson et al., (2013) in their review, in ADHD+ASD samples externalizing behaviors tend to be associated with ADHD (Matsushima et al., 2008), while language and cognitive delays are usually attributed to ASD (Hagberg et al., 2010). Although such a methodology seems reasonable in principle, the behavior of the C+COM and COM groups often appears difficult to compare due to significant differences.

Let us illustrate this point through another example of control-case study that explicitly takes comorbidity into account. There has been some debate about whether autism can be comorbid with developmental language disorders (DLD) - see Tager-Flusberg 2015; Bishop et al., 2016. Starting with the DSM-5, autism and DLD are in fact seen as exclusionary, which implies that

[4] To reiterate a point we already made elsewhere: whenever we use labels such as “ASC+ADHD” or “C+COM”, we do so only for convenience and in reference to current practice, where conditions are usually treated in isolation or in additive combination with one another. In this respect - again - our argument starts from the conservative assumption that conditions may be characterized as identifiable entities which can be studied in isolation or in their interaction.

comorbidity is ruled out as a consequence. Yet, it is commonplace for researchers working on autism and language to talk about two different groups: ASD-LI, and ASD+LI, where LI stands for "linguistic impairment" (Tuller 2021). To complicate things further, ASD+LI groups are usually compared to ASD-LI and DLD groups, because the profile of LI is said to be similar to the profile of DLD. Thus, LI, which is often referred to as "structural language deficits", is *de facto* treated as an additional disorder that autistic people may exhibit. Moreover, LI is often treated as an additional condition that can be factored out by using DLD groups as contrast classes (see Norbury 2005; Tager-Flusberg 2015; Friedman & Sterling 2019).

Now, since Norbury (2005), many researchers have come to share the view that issues with figurative meaning, as well as with some implicatures, are due to structural language problems in ASC, that is, they are attributed to LI comorbidities (Baird & Norbury, 2016). Norbury actually showed that children with DLD performed similarly to LI autistic children in a task where they were asked to choose the most apt metaphor (e.g., "the room was hot. It was: (a) an oven; (b) a blanket; (c) a grill; (d) a spice"). The conclusion has been that not understanding (or "seeing") metaphors relates to poor language abilities. However, similarities or differences between the two relevant groups are not further analyzed. In this case, Norbury and followers have observed difficulties with metaphors without investigating whether such difficulties are actually the same - or at least sufficiently similar - in both groups. Some authors (e.g., Chahboun et al., 2017, Vicente & Falkum, 2021) suggest that they are not: while children with DLD do not understand some metaphors, autistic children understand them *literally*. Moreover, it is a big leap to attribute difficulties with figurative meaning in the ASC+LI profile to the LI component on the basis of shared difficulties observed in DLD individuals, without further analyzing what kind of specific difficulties and what further relevant difficulties each group experiences.

Studies constructed this way are additionally used to predict which traits people with comorbidities will tend to exhibit. For instance, autistic people across the spectrum have been found to have problems with irony and other pragmatic issues that engage the Theory of Mind system (Andrés-Roqueta & Katsos, 2020). Since, as we have just seen, problems with scalar implicatures and figurative meaning are often attributed to comorbid DLD, researchers expect that people in the ASC+LI group will display the characteristic language problems that affect people with DLD *plus* the characteristic problems of ASC (see Andrés-Roqueta & Katsos, 2020 for a recent defense of this view). That is, the method applied in these studies consists in breaking down comorbidities into conditions, assigning traits where they are more likely to belong, and finally adding them up to predict the behavior of people with comorbidities. As we explore below (§3.2), this methodology is probably misguided. In fact, when language difficulties of profile A add up to pragmatic difficulties of profile B, the result is a third profile which shares difficulties with A and with B but also exhibits many further ones, as well as having A's and B's difficulties exacerbated.

To sum up: many researchers in psychiatry aim at discovering what is characteristic of a specific mental condition in order to design successful interventions. Within this paradigm, researchers look for the characteristic features of the relevant condition, that is a cluster of features that would reliably distinguish it from others. Case-control studies are built around the idea that comparisons between groups - i.e., one composed of individuals diagnosed with a given condition and one of individuals without such a diagnosis - may help us reach more general conclusions. This in turn rests on the assumption that members of the relevant categories are sufficiently *homogeneous*. This approach to individual conditions goes hand in hand with an *additive* approach to comorbid conditions. As we explain above, if a person who exhibits two

conditions A and B displays symptom S, which is not characteristic of A, then such symptom is *prima facie* attributed to B. Little research is done about how A and B interact, so it is usually not explored whether the results observed in the A+B group may be due to the interaction between A and B. In the next section we delve deeper into heterogeneity and comorbidity to explore the challenges they raise to this widespread approach.

§3. The challenges of heterogeneity and comorbidity

The approach to mental conditions discussed above runs into problems once heterogeneity and comorbidity are properly taken into account. On the one hand, heterogeneity casts doubts on the idea that subjects who receive the same diagnosis would be sufficiently similar to one another - i.e., contrary to *homogeneity*. On the other hand, comorbidity challenges the compositional view according to which mental conditions can be approached on the basis of a good grasp of conditions themselves and a simple mechanism of addition. By doing so, comorbidity directly challenges *additivity*, but also impacts homogeneity, since we show that comorbidity introduces further heterogeneity (§3.2). We tackle these issues in turn.

§3.1 Heterogeneity (as typically considered)

The issue of heterogeneity in psychiatric classification has been widely acknowledged and has been leveraged as a criticism of the DSM's polythetic structure, which allows for minimally overlapping profiles to qualify for the same diagnosis (Lilienfeld & Treadway 2016). In autism research, heterogeneity appears at various levels of analysis - e.g., genetic, behavioral, phenomenological - and is often listed among the most pressing problems to address (Timimi & McCabe 2016; Lombardo et al., 2019). In what follows we offer an overview of different aspects

of heterogeneity as they are usually discussed within autism research, before introducing a more detailed analysis of comorbidity as a further source of heterogeneity (§3.2).

(i) *Heterogeneous features*: One important source of heterogeneity concerns *features* that are considered diagnostic of a certain condition, which are often left vague enough to allow for different presentations in different individuals. The notion of rigidity or inflexibility in autism is a case in point. Although diagnostic manuals and tools (e.g., DSM-5, ADOS-2) include multiple references to rigid or inflexible features, the notion of rigidity encompasses a wide range of traits such as fixed interests, insistence on sameness, inflexible adherence to routines, black-and-white mentality, intolerance of uncertainty, ritualized patterns of behavior, literalism, and discomfort with change. Such a broad characterization of rigidity allows for significant heterogeneity in terms of phenotypic profiles, given that two individuals can cross the diagnostic threshold for ASD while being rigid or inflexible in very different ways. For instance, someone may be inflexible about rules because of issues with executive functions (e.g., task-switching) whereas someone else may find comfort in following routines.

(ii) *Heterogeneous distribution of features within a condition*: Another source of heterogeneity concerns how features of a given condition are *distributed* in different individuals. Similarly to what happens with (i), individuals who are diagnosed with the same condition end up widely differing from one another. Yet, this is due to different ways of crossing the diagnostic threshold rather than to vagueness of individual features. For example, subjects receive an ASD diagnosis based on two sets of features: socio-communicative problems (which itself comprise a variety of symptoms), and restricted and repetitive behavior or interest, which includes repetitive movements, insistence on sameness, narrow and intense interests, and resistance to change. Since ASD is described through symptoms from apparently unrelated sets, an individual may have

deep socio-communicative problems while exhibiting a moderate degree of insistence on sameness and stereotypies, or any other combination of these sets of features in varying degrees. Although diagnostic tools such as the ADOS-2 set threshold criteria for each feature, the degree to which each feature is exhibited by different individuals may vary greatly. Moreover, as social models also emphasize, individuals also vary with respect to the impact that symptoms have on their lives, and with respect to environmental factors that are hardly comparable across subjects (e.g., degree of social support). Each different distribution of features and degrees to which such features are expressed may therefore give rise to different challenges in the interaction with one's environment.

(iii) *Developmental heterogeneity* indicates that people may greatly differ in their developmental trajectories, with symptoms appearing early in development in some individuals and much later in other cases. Symptoms may be acquired in the course of development, being caused by other symptoms, or rather mature over time. Lack of communicative engagement in autism is an interesting case, as it is unclear whether people who suffer from early regressions (12-30 months of age) are phenotypically similar to people who lack communicative engagement from the beginning (Rogers 2004). Throughout development, core symptoms may even disappear or reduce to such a degree that people who were initially diagnosed as autistic do not meet the DSM-5 characterization at a later stage (Fein et al., 2013; Kelley, Naigles & Fein 2010). Although we cannot expand much further on this point, it is worth noting that developmental trajectories have been discussed among the factors enhancing heterogeneity (Cuthbert 2014), and that autism itself has been characterized as an alternative developmental trajectory of human neural and behavioral development (Johnson 2017). On this latter reading, different degrees of neurodevelopmental atypicalities give rise to multiple developmental routes that bring about an

uneven behavioral profile across individuals, be them neurotypical - e.g., preterm birth - or neurodivergent - e.g., autistic (p. 8).

Heterogeneity raises an important challenge to psychiatric research as normally conducted, as it calls into question the idea that mental conditions are sufficiently homogeneous to justify generalizations in terms of research and intervention. Contrary to the homogeneity assumption, (i)-(iii) show different ways in which members of these (alleged) categories exhibit significantly different symptomatologies and profiles. Even if mental conditions are not conceived as strictly categorical, but rather as prototypical, heterogeneity shows that there are numerous prototypical ways to exemplify a given condition. By contrast, much contemporary research adopting the case-control paradigm assumes that researchers will be working with subjects who belong to *the* prototype of a certain condition. However, heterogeneity shows that such an expectation is unrealistic as at least *several prototypes* would often correspond to a given condition. This issue also has important clinical implications given that individuals who exhibit different profiles tend to react differently to treatment (Feczko et al., 2019). This leaves us in an undesirable situation, where individuals are grouped together in virtue of diagnostic classes that are ultimately heterogeneous. As we show in the next subsection, the situation becomes even more undesirable once we consider the impact of comorbidity as a source of further heterogeneity.

§3.2. Comorbidity as a source of heterogeneity

Rampant comorbidity raises a *prima facie* challenge to the approach currently endorsed by a significant bulk of psychiatric research. Indeed, as we explain above, the structure of the DSM easily allows for the same person to cross the diagnostic threshold for two or more conditions. This makes it harder to identify clear cases and to isolate the characteristic features of each

condition. In this subsection we discuss comorbidity mostly through examples from the autism literature, as autism notoriously exhibits a number of well-known comorbidities to a wide range of conditions including Attention Deficit Hyperactivity Disorder (ADHD), Intellectual disability, Linguistic Impairments, Fragile X-syndrome, and Tourette syndrome (Joshi et al., 2010; Robertson & Eapen 2014). To complicate matters further, the conditions that are comorbid with autism are in turn comorbid with many others. ADHD, for instance, exhibits high rates of comorbidity with Oppositional Defiant Disorder (40%), Conduct Disorder (also 40%), anxiety (34%), and depressive disorders (27%) - see Connor & Doerfler 2008.

We now focus on how comorbidity impacts heterogeneity in terms of features or traits. When it comes to comorbid conditions, it is particularly unclear which features or traits should be ascribed to one or the other condition, or rather to their complex *interaction*. Yet, researchers tend to think additively about comorbidity: for instance, when matching subjects by IQ in a case-control setting, they implicitly assume that IQ would interact in the same way in different individuals with heterogeneous profiles. Here we suggest that interactions of symptoms in comorbid conditions should not be seen as additive but rather as *dynamically complex.*

In order to explain how two or more conditions behave non-additively, we use complex systems as an analogy.[5] A prototypical example of a dynamically complex phenomenon is a hurricane, that is a structure that tends to self-preserve and that is composed of myriads of particles that behave in a peculiar way, clearly constrained by the global pattern or structure of which they are part. As it progresses, the hurricane leaves some of its constituents behind, but replaces them

[5] A note of caution here: we do not want to claim that comorbid conditions *are* complex systems. In fact, there may be relevant disanalogies between the typical complex systems (e.g., far from equilibrium physico-chemical systems, living beings, economical systems, societies, etc.) and systems formed by two or more conditions. However, we think that the analogy is illuminating because it helps us see that interactions between conditions and their characteristic symptoms are more complicated and unpredictable than they are currently taken to be.

with others that it takes from the surrounding matter, causing them to behave in a peculiar way that preserves the global pattern - despite the exchange of matter with the environment surrounding the hurricane. The crucial point is that the behavior of the whole is not explainable, predictable, or deducible from the known behavior of its basic physical properties or "basal conditions" (Kauffman, 1995, Prigogine, 1997). Many philosophers hold that these kinds of systems, which go from a hurricane to a brain, exhibit emergent properties, i.e., new properties that arise from the interaction of known properties under certain boundary conditions (see O'Connor 2021 for an overview). The hallmark of emergentism is unpredictability: even if we knew all there is to be known about the constituents involved in the system, we will not be able to predict the behavior of the system at the global level. In this sense, a complex system is paradigmatically non-additive because the behavior of the whole is not reducible to the known behavior of its constituents when not forming part of the system.[6]

As we explain above, mental conditions are mostly characterized by - and diagnosed on the basis of - *symptoms*, i.e., characteristic ways of behaving or thinking. Autism, for instance, is characterized, among other symptoms, by social and communicative difficulties, repetitive behaviors and restricted interests. Suppose then that a person would be diagnosed both with ASD and ADHD; i.e., the basal conditions, so to speak, would be ASD *and* ADHD. If the interaction of symptoms in comorbid conditions is indeed dynamically complex, the combination of symptoms in these cases will be likely to produce significant effects on the respective symptoms of the individual conditions, thus making such symptoms appear in non-typical ways (where 'typical' refers to 'typical within a given condition').

[6] There is a lively debate in the literature about emergence between those who understand emergence mainly as concerning unpredictability, and so as an epistemic issue, and those who embrace a stronger metaphysical position, according to which unpredictability signals metaphysical novelty (see Vicente, 2013). We think we can sidestep this debate, as our point is that knowledge of putative basal conditions of individuals is not sufficient to describe such individuals. That is, we do not need to commit to metaphysical emergence for our argument to work.

One of such effects that we deem significant is *intensification*. For instance, people with ASD+ADHD do not simply exhibit the socio-affective and communicative problems typical of ASD *plus* the hyperactivity, impulsivity, and inhibition-related problems typical of ADHD. Rather, the co-occurence of both conditions has been found to be correlated with the appearance of typical features of each condition at a higher degree of severity (Reiersen et al., 2007). Thus, in a study on children with ASD, ADHD, and ASD+ADHD (age 6-13), Yerys and colleagues (2009) found that children with ASD+ADHD exhibited more severe autistic traits (especially along the social and verbal dimension), as well as more severe attentional issues (inattention, hyperactivity, and impulsivity symptoms). Similar results have been reported in the case of ADHD and learning disabilities (LD). Children with ADHD+LD have been found to experience more severe academic difficulties *and* more severe behavioral problems compared to children with ADHD and LD considered individually (Smith & Adams 2006). Individual features may thus be expressed at a higher degree of severity depending on how comorbid conditions interact. Such intensification effects are the bread and butter of far-from-equilibrium complex systems, to the extent that in his first work on the topic, Prigogine (1955) observed that far from equilibrium chemical systems exhibited reaction rates that are not even considered in equilibrium thermodynamics of homogeneous systems.

More generally, how a symptom is expressed depends on contextual factors and, *inter alia*, on what other conditions the person exhibits. This is, again, similar to what we observe in complex systems: constituents become part of a dynamics that alters their typical causal profiles. For instance, complex dynamics often involve the elimination of degrees of freedom of the whole system's constituents, such that constituents are constrained in their behavior and only exhibit some among the possible ways in which they *could* behave (see Wilson 2010 on the notion of

elimination of degrees of freedom). This is something that can also occur when different symptoms appear together: the manifestation of some symptoms may modulate the usual manifestation of some other symptoms. *Weakening* effects of some condition over the expression of another (or at least of some of its characteristic symptoms) are also a possibility, although such cases might not qualify as comorbid from a DSM perspective. Indeed, the weakening effects might make it so that the symptoms of a condition are not observed as being prominent enough to cross the diagnostic threshold for the said condition. In such cases, complex dynamics among constituents might actually *reduce* comorbidity (with respect to the basal conditions), although the same would not apply to heterogeneity, given that they would give rise to yet another profile of individuals who exhibit attenuated symptoms.

Another significant effect that we observe in complex systems concerns the appearance of *novel features* that are not characteristic of basal conditions taken individually, but which emerge once the two conditions are simultaneously in place. In situations of comorbidity we observe a similarly complex dynamic. In some cases, the interaction between symptoms allegedly pertaining to different conditions causes the individual to exhibit new conditions altogether. For instance, individuals exhibiting ASD+ADHD often suffer from anxiety and mood disorders (see Gordon-Lipkin et al., 2018 for a recent study on children). Similarly, higher levels of atypical behaviors (e.g., atypical eating behavior, abnormal sleep patterns, temper tantrums, and self-injurious behavior) have been observed in autistic people with low levels of expressive language and nonverbal IQ, when compared to people with low levels of expressive language and nonverbal IQ or to autistic people with higher levels of expressive language and IQ (Dominick et al., 2007). Another illustrative case in this respect is the combination of intellectual disability and autism, which is associated with a broad range of concomitant health issues and psychiatric

disorders (Bertelli & Bianco 2021). While autistic people can learn on their own, intellectually disabled people can learn from others; however, the combination of ASD and ID deeply hinders these two ways of learning (Jordan 2013). In this case, we may be tempted to see the learning difficulties of ASD+ID individuals as additive, since they exhibit both difficulties related to ID *and* difficulties related to ASD. However, the combination of both seems to result in an additional symptom, i.e., non-verbality (which affects 25-30% of autistic people). Non-verbality is undoubtedly a larger issue, and it is premature to reduce it to the addition of ID's and ASD's learning difficulties. Yet, it is important to notice that lack of language has widespread effects on all sorts of dimensions, including more severe learning difficulties (since a non-verbal person cannot learn from verbal means). Beyond such learning difficulties, non-verbality in the context of autism may have an impact on cognition (Hinzen et al., 2019), as well as on detachment and anxiety levels, given that non-verbal or minimally-verbal people often strive, but fail, to communicate with others (Dominick et al., 2007). The extreme levels of anxiety that non-verbal people experience have an impact on their learning abilities as well, making interventions particularly difficult. In this case, symptoms tend to form a self-sustaining network that is emergent over the typical symptoms of ASD and of ID alone (more on this in §4). Self-sustaining dynamics are also a hallmark of complex systems, as illustrated by the hurricane example.

Although we do not want to commit to interactions between symptoms *actually* forming complex systems, we regard the analogy with physical complex systems as illuminating, especially to understand our claim that symptoms do not behave additively. The main idea is that interaction among symptoms or features changes the nature of symptoms themselves and gives rise to emergent dynamics. Once again, we are taking a conservative stance here, since we are

only considering the interaction between typical symptoms of DSM conditions, without looking at the wider context of interaction that comprises personality traits, environmental factors and material, social and ideological structures or constraints. Yet, even within such a broader perspective, complex combinations of features are likely to behave non-additively and unpredictably. For instance, rates of depression and anxiety in the autistic population appear to be inversely proportional to symptom severity and directly proportional to verbal IQ (Mazurek & Kanne, 2010; but see, Strang et al., 2012). If this is correct, autistic people with higher verbal IQ *and* less severe symptoms - i.e., two features that would be regarded as advantageous in many contexts - would be more likely to develop depressive disorders. Although this point is quite speculative given that little is known about the relevant causal underpinnings, we want to stress that complex systems dynamics may also affect the interaction between individual characteristics and environmental features (including lack of support, etc.). In this respect, for instance, it is likely that the higher rates of depression observed in highly verbal autistic people may relate to our societies lacking inclusiveness.

Taking stock: the discussion prompted by comorbidity forces us to call the assumption of additivity into question. By overlooking complexity and the effects it engenders, researchers obtain an overly narrow picture of the groups of individuals under study. Indeed, by focusing almost exclusively on basal conditions and their addition, current research risks missing some significant effects emerging by their interactions - such as intensification or novel features as discussed above. Given this level of complexity, it simply does not seem feasible to predict and capture the issues related to a compound condition starting from its components. Overlooking the complexity surrounding the (multiple) emergent features in cases of comorbidity thus implies missing out on significant aspects of these conditions (be them clinical or subclinical). This has

important implications for intervention, as - for instance - evidence-based interventions targeted toward autistic people will not be *ipso facto* applicable to autistic people with comorbidities. Interventions in the case of people with comorbidities cannot be additive either, as they cannot work as a mere combination of interventions targeted toward individual conditions. Evidence-based interventions for cases of comorbidity must rather arise from more accurate descriptions of the relevant groups, which should be explored and addressed directly.

In the next section we offer some suggestions on how to tackle the challenge raised by comorbidity and its impact on heterogeneity. We first explore a group of approaches that propose various forms of *transdiagnostic stratification* and we explain why they help us move some preliminary steps in the right direction (Cuthbert 2014; Lombardo et al., 2019). We then briefly discuss network models, which have focused on comorbidity more explicitly (Cramer et al., 2010; Fried et al., 2017).

§4. Sketching an alternative

In the previous sections we discussed two assumptions - i.e., homogeneity and additive comorbidity - which are still widespread in psychiatric research despite the criticisms surrounding the DSM model. We then showed that facts about comorbidity constitute a serious challenge to these assumptions, since comorbidity exacerbates heterogeneity by giving rise to new conditions and developmental as well as phenotypic profiles, thus pressing us forcefully to move beyond traditional classifications. Here we focus on some alternative approaches that have been gaining traction in the wake of the discontent with the DSM. We are particularly interested in exploring how they fare with respect to what we take to be the relevant features of comorbidity - i.e., the fact that individuals rarely exhibit just one condition, that characteristics of

different conditions interact in complex ways by giving rise to intensification, weakening effects, and novel symptoms, and that similar symptomatologies may relate to different basal conditions.

As we mention above, the first obvious candidate for an alternative view is the *Research Domain Criteria* project (RDoC), which aims to provide a novel framework for research in psychopathology by moving away from a categorical and descriptive approach and by rooting classification in circuit-based and behavioral dimensions (Cuthbert 2014). If successful, the program would allow clinicians to individuate particular *subtypes* or critical locations along the relevant dimensions, marking a promising move from case-control settings to studies focused on dimensions of functioning and their distribution across clinical and nonclinical populations. In this respect, the RDoC aims to move psychiatry closer to precision medicine or stratification (Lilienfeld & Treadway 2016; Tabb 2020).

Generally speaking, the goal of stratification approaches within medical research is to identify meaningful substructures within a complex sample through supervised or unsupervised methods based on data-mining, where different sources of information converge to identify significant substructures or clusters. To be reliable, these approaches require a certain degree of *breadth* (i.e., large sample sizes) and *depth* (i.e., multiple levels of data available for each individual), to which we may add *precision* (i.e., data have to be precise as opposed to vague) - see Lombardo et al., 2019; Tuller 2021 for some examples. Ideally, stratification would end up forming groups that are more or less homogeneous with respect to genotypic and phenotypic features, as well as in terms of treatment response. For instance, several research groups have recently attempted to identify biologically, cognitively, and behaviorally meaningful subtypes for autism (see Wolfers et al., 2019 for a review). In these studies, stratification sets out to identify a variety of "signatures" or "markers" that may be indicative of current or future autism development.

Notably, stratification is often presented by researchers as an effective approach for tackling *heterogeneity* (Lombardo et al., 2019). Indeed, stratification aims to describe individuals more accurately by generating clusters of symptoms at different levels (i.e., biological, cognitive, behavioral, etc.) out of large samples of data. For instance, we may think about stratification approaches as being able to differentiate between different subtypes of autism that exhibit different profiles, responses to treatment, etc. Such an approach is in principle agnostic with respect to comorbidities, as these can be conceived as lying outside the features of conditions or they can be included as constitutive of heterogeneity. In the former case, an individual may exhibit certain characteristic symptoms of a homogeneous condition, and thus unequivocally belong to one cluster or subgroup. Yet, such a condition may still be heterogeneous in a different sense, because it may include people with very different profiles due to comorbidities that affect how symptoms are expressed in complex ways (e.g., intensification, weakening, appearance of novel features). By not attending to comorbidities carefully, stratification approaches thus risk missing out on the explanation for why symptoms are expressed in the way they are (or even for why they are expressed at all).

In sum: given how features complexly interact with each other, narrowing down the focus to the characteristic features of a certain condition risks resulting in uninformative clusters. We take this to show that stratification approaches need to broaden their perspective to include comorbidity as an additional source of heterogeneity. One of the purposes of stratification models is to account for the causal mechanisms that make some clusters of features appear together, so it is crucial to include features that have an effect on other features.

One possible way to do this would be to commit more radically to a *transdiagnostic* approach that cuts across current diagnostic boundaries and specifically targets comorbidity (actually,

Lombardo et al., 2019 themselves express sympathy for a transdiagnostic approach). Transdiagnostic research proposes a departure from condition-based approaches and aims at identifying the common temperamental, psychological, cognitive, emotional, interpersonal, and behavioral processes that underpin a broad array of diagnostic presentations (McGorry et al., 2018; Krueger & Eaton 2015). The most comprehensive review to date concludes that, despite the formal commitment to such an approach, the majority of self-labeled transdiagnostic studies are at odds with the paradigm as they investigate either individual symptoms or a single disorder (Fusar-Poli et al., 2019). Moreover, current transdiagnostic research is grounded on a prototype-based approach to mental conditions, where prototypes are usually built around the characteristic features of a given condition. However, given rampant heterogeneity and comorbidity, these prototypes would only minimally represent individuals who exhibit the relevant condition(s). To better implement a transdiagnostic strategy, it would be advisable to form clusters based on fine-grained features across levels, independently of whether such features are characteristic of one condition or another. In other words, it would be irrelevant to establish from the outset whether a given feature belongs to condition A or B (comorbid with A). The main issue in building a stratified similarity space would be to see whether a given region in the space, where several symptoms overlap, actually describes a group of individuals. Once we are able to build a similarity space with these finer-grained and multi-level features, we can start asking more meaningful questions about causation: e.g., what kind of explanatory networks their features form, what kind of ultimate causes sustain the network, etc.

Existing stratification approaches also encounter other challenges. For instance, proponents of the RDoC project, who are quite explicit about embracing a transdiagnostic view (Cuthbert 2014), have been criticized for their undue emphasis on neurobiological components at the

expense of subjective and behavioral manifestations (Lilienfeld & Treadway 2016).[7] Transdiagnostic stratification approaches should thus be complemented by other models that focus more explicitly on the psychological level. In our view, while there is of course a role for mechanistic explanations, which would be able to explain psychological manifestations in terms of neurobiological mechanisms, explanations at the purely psychological level are also necessary to disentangle the complex ways in which symptoms interact and how such interaction unfolds over time.

In this respect, network models of psychopathology (Borsboom 2017; Borsboom et al., 2019) appear to be well-suited to address these issues. These accounts set out to debunk the very idea of underlying disruptions causing symptoms - i.e., traditional latent models - and regard mental disorders as emerging from the complex interaction among symptoms. Some strands of network models have also focused on *comorbidity* more specifically (Cramer et al., 2010; Fried et al., 2017), with a special emphasis on co-occurring clusters of symptoms forming stable networks - e.g., depression and anxiety. Given their focus on symptoms as the relevant unit of analysis (Borsboom 2017), network models are bound to be particularly useful to researchers in the preliminary process towards understanding patterns, configurations, and developmental trajectories at the personal level (see for instance Deserno et al., 2018 on autism). In this sense they could fruitfully supplement the stratification approaches mentioned above.

However, it is worth stressing that classic formulations of network models risk being too conservative in their focus on surface-level description of symptoms and their relation to one

[7] Other transdiagnostic stratification projects which we cannot discuss in detail here - such as ROAMER (Schumann et al. 2014) and HiTOP (Kotov et al. 2017) - defend a broader view in this respect and attempt to complement the RDoC framework by adding constructs and dimensions that take into account more refined behavioral and developmental characterizations.

another within a condition (e.g., major depression). In our view, the observation of how symptoms interact with one another to form self-sustaining networks should not foreshadow the etiological component, or at the very least the commitment to the idea that networks emerge *for a reason* (be it genetic makeup, developmental trajectory, life experiences, personality traits, disrupted relation with the environment, and so on). Identifying complex interactions between symptoms could therefore work as a promising starting point for an alternative approach towards mental conditions, but should not be understood as the be-all and end-all of such an enterprise (see Haig & Vertue 2010 for a similar point). Fortunately, some versions of network models acknowledge these points. For instance, Cramer and colleagues (2010) approach comorbidity by explicitly gesturing towards dynamical systems to account for complex interactions between symptoms *across* conditions. Others have *de facto* relinquished the assumption that DSM-like symptoms should be taken as the relevant unit of analysis to grant that other variables - e.g., neurobiological makeup, life events, socio-cultural context - are likely to play a significant role in a network's formation and development (Fried & Cramer 2017). Such a broader outlook with respect to comorbidity and etiology, combined with an approach that acknowledges multiple levels of analysis and causation, make these strands of network models particularly suited to the construction of transdiagnostic clusters that more accurately describe groups of individuals.

Taking stock: the search for satisfactory alternatives to the current model based on DSM classifications has brought to the fore interesting aspects of transdiagnostic stratification approaches and network models. While the former have the merit of cutting across existing categories to identify the relevant underpinnings of different conditions, the latter provide us with a promising strategy to observe and describe complex interactions at the level of symptoms.

We therefore regard the two approaches as possibly complementing each other to address the challenges discussed above.

With respect to our case study, we have reasons to think that such a hybrid approach would be particularly helpful if applied to people with an autism diagnosis. On the one hand, autistic symptoms as described by the DSM are assumed to be effects of a neurodevelopmental condition, which is still far from being fully understood in its genetic and neurobiological underpinnings. On the other hand, even the characteristic cognitive and behavioral symptoms of autism appear disunified, as the DSM description comprises two areas with no obvious connection with one another, i.e., social communication and interaction, and restricted and repetitive behaviors. In this respect, autism showcases the obvious need for etiological accounts that bring together cognitive-behavioral symptoms, (epi)genetic factors, and neurodevelopmental trajectories. Yet, given the staggering rate of comorbidity affecting people diagnosed as autistic, to achieve such a goal we require models of complex interactions between symptoms at multiple levels (executive functions, language development, anxiety, activity levels, intellectual abilities of different sorts, sensory integration, adaptability, self-consciousness, etc.). In particular, such models could shed some light on developmental heterogeneity (see §3.1), which appears particularly puzzling in autism and includes a broad spectrum of phenomena ranging from regressive pathways to loss of autism diagnosis. The lack of a standard developmental trajectory in people with an autism diagnosis makes it particularly difficult for clinicians, families, and caregivers to understand and predict how children with a diagnosis are likely to develop. Being able to identify clusters of developmental pathways, while at the same time enhancing our understanding of how different factors interact with one another (from neurobiological conditions to co-occurring symptoms putatively ascribed to other DSM conditions), would at

least provide us with some preliminary understanding of why individuals with an autism diagnosis exhibit such a wide range of developmental paths.

Conclusion

The approach we describe in sections §1 and §2 is firmly rooted in psychiatric research. Once we acknowledge that heterogeneity exacerbated by comorbidity is pervasive, we realize that the current approach is ineffective as it creates the illusion that we can pull conditions apart and still deepen our understanding of people exhibiting such a condition. To recap our argument: even if we assume that DSM categories describe conditions pertaining to individuals, the likelihood that such categories would be useful to know more about individuals is small. First, DSM conditions themselves allow for significant heterogeneity. Second, and more importantly, DSM conditions rarely appear alone. Such rampant comorbidity gives rise to complex interactions between symptoms supposedly belonging to different conditions, thereby generating intensification, emergent, and maybe even weakening effects. In turn, such effects render knowledge of "pure" conditions idle. In sum, even starting from conservative premises and granting some core assumptions to DSM classifications, we reach conclusions similar to those who distrust the DSM approach.

On a more positive note, we think that stratification approaches represent a good starting point to create transdiagnostic clusters that will take into account complex and multi-layered profiles as well as networks of symptoms. This might lead to a thorough revision of psychiatric nosology, one in which we dispense with the notion of comorbidity (and of condition) altogether. This way, psychiatric research would move from dealing with idealized *mental conditions* to identifying meaningful groups of *individuals* exhibiting clinically relevant features. We take it that this view

is compatible (at least in some cases) with seeing such clinically relevant features as a result of a specific way of coupling a neurodivergent individual way of life with specific ways of enforcing neurotypical standards of interaction.

References

Allsopp, K., Read, J., Corcoran, R., & Kinderman, P. (2019). Heterogeneity in psychiatric diagnostic classification. *Psychiatry research*, 279, 15-22.

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

Andrés-Roqueta, C. & Katsos, N. (2020). A distinction between linguistic- and social-pragmatics helps the precise characterisation of pragmatic challenges in children with Autism Spectrum Disorders and Developmental Language Disorder. *Journal of Speech, Language and Hearing Research.*

Aragona, M.

(2009a). About and beyond comorbidity: does the crisis of the DSM bring on a radical rethinking of descriptive psychopathology? *Philosophy, Psychiatry, & Psychology*, 16(1), 29-33.

(2009b). The role of comorbidity in the crisis of the current psychiatric classification system. *Philosophy, Psychiatry, & Psychology*, *16*(1), 1-11.

Baird, G., & Norbury, C. F. (2016). Social (pragmatic) communication disorders and autism spectrum disorder. *Archives of Disease in Childhood*, 101(8), 745-751.

Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *Journal of child Psychology and Psychiatry*, *30*(2), 285-297.

Bertelli, M. O., & Bianco, A. (2021). Assessment of Psychiatric Disorders. In *Mental Health, Intellectual and Developmental Disabilities and the Aging Process* (pp. 45-60). Springer.

Bervoets, J., & Hens, K. (2020). Going Beyond the Catch-22 of Autism Diagnosis and Research. The Moral Implications of (Not) Asking "What Is Autism?". *Frontiers in Psychology*, *11*.

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & Catalise Consortium. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLOS one*, 11(7), e0158753.

Blashfield, R. K., Keeley, J. W., Flanagan, E. H., & Miles, S. R. (2014). The cycle of classification: DSM-I through DSM-5. *Annual review of clinical psychology*, *10*, 25-51.

Borsboom, D., Cramer, A. O., & Kalis, A. (2019). Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences*, *42*.

Borsboom, D. (2017). A network theory of mental disorders. *World psychiatry*, *16*(1), 5-13.

Botha, M., Dibb, B., & Frost, D. M. (2020). "Autism is me": an investigation of how autistic individuals make sense of autism and stigma. *Disability & Society*, 1-27.

Brewer, N., Young, R. L., & Barnett, E. (2017). Measuring theory of mind in adults with autism spectrum disorder. *Journal of autism and developmental disorders*, *47*(7), 1927-1941.

Chahboun, S., Vulchanov, V., Saldana, D., Eshuis, H., & Vulchanova, M. (2017). Can you tell it by the prime? A study of metaphorical priming in high-functioning autism in comparison with

matched controls. *International journal of language & communication disorders*, *52*(6), 766-785.

Chapman, R.

(2020). The reality of autism: On the metaphysics of disorder and diversity. *Philosophical Psychology*, *33*(6), 799-819.

(2019). Neurodiversity theory and its discontents: Autism, schizophrenia, and the social model of disability. *The Bloomsbury Companion to Philosophy of Psychiatry*. Ş. Tekin & R. Bluhm (eds.), 371-90.

Chevallier, C., Noveck, I., Happé, F., & Wilson, D. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia*, *49*(3), 507-517.

Connor, D. F., & Doerfler, L. A. (2008). ADHD with comorbid oppositional defiant disorder or conduct disorder: discrete or nondistinct disruptive behavior disorders? *Journal of Attention Disorders*, 12(2), 126-134.

Cooper, R.

(2020). The concept of disorder revisited: Robustly value-laden despite change. In *Aristotelian Society Supplementary Volume* 94(1), 141-161. Oxford University Press.

(2014). *Diagnosing the diagnostic and statistical manual of mental disorders*. Routledge.

(2005). *Classifying madness*. Springer.

Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry research*, 166(2-3), 210-222.

Cramer, A. O., Waldorp, L. J., Van Der Maas, H. L., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and brain sciences*, *33*(2-3), 137-150.

Cuthbert, B. N. (2014). The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry*, *13*(1), 28-35.

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC medicine*, 11(1), 1-8.

Deserno, M. K., Borsboom, D., Begeer, S., & Geurts, H. M. (2018). Relating ASD symptoms to well-being: Moving across different construct levels. *Psychological Medicine*, *48*(7), 1179-1189.

Dominick, K. C., Davis, N. O., Lainhart, J., Tager-Flusberg, H., Folstein, S. (2007) Atypical behaviors in children with autism and children with a history of language impairment. *Research on Developmental Disabilities*, 28:145-162.

Feczko E, Miranda-Dominguez O, Marr M, Graham AM, Nigg JT, Fair DA (2019): The heterogeneity problem: Approaches to identify psychiatric subtypes. *Trends in Cognitive Science* 23:584–601.

Fein, D., Barton, M., Eigsti, I. M., Kelley, E., Naigles, L., Schultz, R. T., Stevens, M., Helt, M., Orinstein, A., Rosenthal, M., Troyb, E., & Tyson, K. (2013). Optimal outcome in individuals

with a history of autism. *Journal of child psychology and psychiatry, and allied disciplines*, *54*(2), 195–205. https://doi.org/10.1111/jcpp.12037

Feinstein, A. R. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of chronic diseases*, 23(7), 455-468.

Fellowes, S. (2021). The Value of Categorical, Polythetic Psychiatric Diagnoses. *British Journal for the Philosophy of Science*.

First, M. B. (2012). The National Institute of Mental Health Research Domain Criteria (RDoC) project: moving towards a neuroscience-based diagnostic classification in psychiatry. *Philosophical issues in psychiatry II: Nosology*, 12, 18.

Fried, E. I., van Borkulo, C. D., Cramer, A. O., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1-10.

Fried, E. I., & Cramer, A. O. (2017). Moving forward: challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science*, *12*(6), 999-1020.

Friedman, L., & Sterling, A. (2019). A review of language, executive function, and intervention in autism spectrum disorder. In *Seminars in speech and language*. 40(4), 291-304). Thieme Medical Publishers.

Fusar-Poli, P., Solmi, M., Brondino, N., Davies, C., Chae, C., Politi, P., ... & McGuire, P. (2019). Transdiagnostic psychiatry: a systematic review. *World Psychiatry*, 18(2), 192-207.

Giora, R., Gazal, O., Goldstein, I., Fein, O., & Stringaris, A. (2012). Salience and context: Interpretation of metaphorical and literal language by young adults diagnosed with Asperger's syndrome. *Metaphor and Symbol*, *27*(1), 22-54.

Gordon-Lipkin, E., Marvin, A. R., Law, J. K., & Lipkin, P. H. (2018). Anxiety and mood disorder in children with autism spectrum disorder and ADHD. *Pediatrics*, 141(4).

Haig, B. D., & Vertue, F. M. (2010). Extending the network perspective on comorbidity. *Behavioral and Brain Sciences*, *33*(2-3), 158.

Haslam, N. (2014). Natural kinds in psychiatry: Conceptually implausible, empirically questionable, and stigmatizing. In *Classifying psychopathology: Mental kinds and natural kinds*, 11-28.

Hinzen, W., Slušná, D., Schroeder, K., Sevilla, G., & Vila Borrellas, E. (2019). Mind–Language=? The significance of non-verbal autism. *Mind & Language*.

Hagberg, B. S., Miniscalco, C., & Gillberg, C. (2010). Clinic attenders with autism or attention-deficit/hyperactivity disorder: cognitive profile at school age and its relationship to preschool indicators of language delay. *Research in developmental disabilities*, *31*(1), 1-8.

Hyman, S. E. (2010). The diagnosis of mental disorders: the problem of reification. *Annual review of clinical psychology*, 6, 155-179.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... & Wang, P. (2010). *Research domain criteria (RDoC): toward a new classification framework for research on mental disorders*. American Psychiatric Association.

Johnson, M. H. (2017). Autism as an adaptive common variant pathway for human brain development. *Developmental Cognitive Neuroscience, 25*, 5-11.

Jordan, R. (2013) (2nd Edition) *Autism with Severe Learning Difficulties*. Souvenir Press.

Joshi G., Petty C., Wozniak J., Henin, A., Fried, R., Galdo, M., Kotarski, M., Walls, S., Biedermanet, J. (2010) The heavy burden of psychiatric comorbidity in youth with autism spectrum disorders: a large comparative study of a psychiatrically referred population. *Journal of Autism and Developmental Disorders* 40:1361-1370.

Kapp, S. K., & Ne'eman, A. (2020). Lobbying autism's diagnostic revision in the DSM-5. In *Autistic Community and the Neurodiversity Movement* (pp. 167-194). Palgrave Macmillan.

Kasirer, A., & Mashal, N. (2016). Comprehension and generation of metaphors by children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 32*, 53-63.

Kauffman, S. (1995) *At home in the universe.* New York: Oxford University Press.

Kelley, E., Naigles, L., & Fein, D. (2010). An in-depth examination of optimal outcome children with a history of autism spectrum disorders. *Research in autism spectrum disorders*, *4*(3), 526-538.

Kendell, R., & Jablensky, A. (2003). Distinguishing between the validity and utility of psychiatric diagnoses. *American journal of psychiatry*, 160(1), 4-12.

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, 20(4), 442-462.

Kim, J. (2006). Emergence: Core ideas and issues. *Synthese* 151, 547–559.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., ... & Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of abnormal psychology*, *126*(4), 454.

Krueger, R. F., & Eaton, N. R. (2015). Transdiagnostic factors of mental disorders. *World Psychiatry*, *14*(1), 27.

Kupfer, D. J., First, M. B., & Regier, D. A. (Eds.). (2002). *A research agenda for DSM V*. American Psychiatric Pub.

Lenzenweger, M.F. (2010). Contemplations on Meehl (1986). *Contemporary directions in psychopathology*, 187-204.

Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual review of clinical psychology*, *12*, 435-463.

Lombardo, M. V., Lai, M. C., & Baron-Cohen, S. (2019). Big data approaches to decomposing heterogeneity across the autism spectrum. *Molecular psychiatry*, 24(10), 1435-1450.

McGorry, P. D., Hartmann, J. A., Spooner, R., & Nelson, B. (2018). Beyond the "at risk mental state" concept: transitioning to transdiagnostic psychiatry. *World Psychiatry*, *17*(2), 133-142.

Maj, M.

(2011). When does depression become a mental disorder? *The British Journal of Psychiatry*, 199, 85–86.

(2005). ‘Psychiatric comorbidity’: an artifact of current diagnostic systems? *The British Journal of Psychiatry*, 186(3), 182-184.

Matson, J. L., Rieske, R. D., & Williams, L. W. (2013). The relationship between autism spectrum disorders and attention-deficit/hyperactivity disorder: An overview. *Research in developmental Disabilities*, 34(9), 2475-2484.

Matsushima, N., Miyawaki, D., Tsuji, H., Takahashi, K., Horino, A., Kawaguchi, T., ... & Kiriike, N. (2008). Evaluation of attention-deficit/hyperactivity disorder symptoms in male children with high-functioning pervasive developmental disorders. *Osaka City Medical Journal*, *54*(1), 1-10.

Maung, H. H. (2016). Diagnosis and causal explanation in psychiatry. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *60*, 15-24.

Mazurek, M.O. & Kanne, S.M. (2010). Friendship and internalizing symptoms among children and adolescents with ASD. *Journal of Autism and Developmental Disabilities*, 24: 438-49.

Milton, D. (2016). Re-thinking autism: diagnosis, identity and equality: Re-thinking autism: diagnosis, identity and equality, edited by Katherine Runswick-Cole, Rebecca Mallet and Sami Timimi, London, Jessica Kingsley Publishers, 2016, 336 pp.,£ 18.99 (paperback), ISBN 978-1-78-450027-6.

Morsanyi, K., Stamenkoviæ, D., & Holyoak, K. J.

(2020). Metaphor processing in autism: A systematic review and meta-analysis. *Developmental Review, 57*, 100925.

(2019). Analogical reasoning in autism. *Thinking, Reasoning, and Decision Making in Autism*. Routledge.

Murphy, D. (2006). *Psychiatry in the scientific image*. MIT Press.

Norbury, C. F. (2005). The relationship between Theory of Mind and metaphor: Evidence from children with language impairment and autistic spectrum disorder. *British Journal of Developmental Psychology*, 23, 383-399.

O'Connor, T. (2021). Emergent Properties. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.) https://plato.stanford.edu/archives/win2021/entries/properties-emergent

Olbert, C. M., Gala, G. J., & Tupler, L. A. (2014). Quantifying heterogeneity attributable to polythetic diagnostic criteria: theoretical framework and empirical application. *Journal of Abnormal Psychology*, 123(2), 452.

Pijnacker, J., Hagoort, P., Buitelaar, J., Teunisse, J.-P., & Geurts, B. (2009). Pragmatic inferences in high-functioning adults with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, 39(4), 607–618.

Pinkus, H. A., Tew, J. D., & First, M. B. (2004). Psychiatric comorbidity: is more or less. *World Psychiatry*, 3, 24-25.

Prigogine, I.

(1996). *The End of Certainty*. The Free Press.

(1955). *Introduction to Thermodynamics of Irreversible Processes*. Charles C. Thomas Publisher.

Regier, D. A., Narrow, W. E., Kuhl, E. A., Kupfer, D. J., & American Psychopathological Association. (2010). *The conceptual evolution of DSM-5*. American Psychiatric Pub.

Reiersen, A. M., Constantino, J. N., Volk, H. E., & Todd, R. D. (2007). Autistic traits in a population-based ADHD twin sample. *Journal of Child Psychology and Psychiatry*, 48(5), 464-472.

Robertson, M. M., & Eapen, V. (2014). Tourette's: syndrome, disorder or spectrum? Classificatory challenges and an appraisal of the DSM criteria. *Asian Journal of Psychiatry*, *11*, 106-113.

Rogers, S. J. (2004). Developmental regression in autism spectrum disorders. *Mental retardation and developmental disabilities research reviews* 10(2), 139-143.

Samuels, R. (2009). Delusion as a natural kind. Psychiatry as cognitive neuroscience: *Philosophical perspectives*: 49-79.

Schumann, G., Binder, E. B., Holte, A., de Kloet, E. R., Oedegaard, K. J., Robbins, T. W., ... & Ciccocioppo, R. (2014). Stratified medicine for mental disorders. *European Neuropsychopharmacology*, 24(1), 5-50.

Smith, T., & Adams, G. (2006). The effect of comorbid AD/HD and learning disabilities on parent-reported behavioral and academic outcomes of children. *Learning Disability Quarterly*, 29(2), 101-112.

Strang, J.F., Kenworthy, L., Daniolos, P., Case, L., Wills, M.C., Martin, A., et al. (2012). Depression and anxiety symptoms in children and adolescents with autism spectrum disorders without intellectual disability. *Research in Autism Spectrum Disorders*, 6:406–412

Tabb, K. (2020). Should Psychiatry Be Precise? Reduction, Big Data, and Nosological Revision in Mental Health Research. *Levels of Analysis in Psychopathology*. Edited by Kenneth S. Kendler, Josef Parnas, Peter Zachar. Cambridge University Press.

Tager-Flusberg, H. (2015). Defining language impairments in a subgroup of children with autism spectrum disorder. *Science China Life Sciences*, 58(10), 1044-1052.

Timimi, S., McCabe, B., Runswick-Cole, K., & Mallett, R. (2016). What have we learned from the science of autism. *Re-thinking autism: Diagnosis, identity and equality*, 30-48.

Tsou, J. (2016). Natural kinds, psychiatric classification and the history of the DSM. *History of Psychiatry*, *27*(4), 406-424.

Tuller, L. (2021). Autism and language modularity. Plenary talk at The Romance Turn conference.

van Loo, H. M., & Romeijn, J. W. (2015). Psychiatric comorbidity: fact or artifact? *Theoretical Medicine and Bioethics*, 36(1), 41-60.

Vella, G., Aragona, M., & Alliani, D. (2000). The complexity of psychiatric comorbidity: a conceptual and methodological discussion. *Psychopathology*, 33(1), 25-30.

Verhoeff, B. (2013). Autism in flux: a history of the concept from Leo Kanner to DSM-5. *History of psychiatry*, *24*(4), 442-458.

Vicente A. (2013) Where to Look for Emergent Properties. *International Studies in the Philosophy of Science*, 27:2, 137-156, DOI: 10.1080/02698595.2013.813256

Vicente, A., & Falkum, I. L. (2021). Accounting for the preference for literal meanings in ASC. *Mind and Language*.

Vulchanova, M., Saldaña, D., Chahboun, S., & Vulchanov, V. (2015). Figurative language processing in atypical populations: The ASD perspective. *Frontiers in human neuroscience*, *9*, 24.

Walenski, M., & Love, T. (2017). The real-time comprehension of idioms by typical children, children with specific language impairment and children with Autism. *Journal of speech pathology & therapy*, *3*(1).

Widiger, T. A., & Crego, C. (2015). Process and content of DSM-5. *Psychopathology Review*, *2*(1), 162-176.

Wilson, J. (2010). Non-reductive physicalism and degrees of freedom. *The British Journal for the Philosophy of Science*, 61, 279–311

Wolfers, T., Floris, D. L., Dinga, R., van Rooij, D., Isakoglou, C., Kia, S. M., ... & Peng, H. (2019). From pattern classification to stratification: towards conceptualizing the heterogeneity of Autism Spectrum Disorder. *Neuroscience & Biobehavioral Reviews.*

Yang, J., Zhou, S., Yao, S., Su, L., & McWhinnie, C. (2009). The relationship between theory of mind and executive function in a sample of children from mainland China. *Child psychiatry and human development*, 40(2), 169-182.

Yerys, B. E., Wallace, G. L., Sokoloff, J. L., Shook, D. A., James, J. D., and Kenworthy, L. (2009). Attention deficit/hyperactivity disorder symptoms moderate cognition and behavior in children with autism spectrum disorders. *Autism Research* 2, 322–333.

Zachar, P. (2000). Psychiatric disorders are not natural kinds. *Philosophy, Psychiatry, & Psychology*, 7(3), 167-18.