Gloomy Prospects and Roller Coasters: Finding Coherence in Genome Wide Association Studies

Carl F. Craver, Mikhail Dozmorov, Mark Reimers, and Kenneth S. Kendler*†

We address Turkheimer's argument that genome-wide association studies of behaviors and psychiatric traits will fail to produce coherent explanations. We distinguish two major sources of potential incoherence (heterogeneity and complexity), showing how they independently threaten coherence and considering how they are being and might better be addressed with protein and other databases and network tools.

1. Two Coherence Problems. Some people like roller coasters; some do not. If we study this trait across generations, we find it is heritable (as required by Turkheimer's first law of behavioral genetics: "All human behavioral traits are heritable [to some extent]"; 20[illegible]XXX). This heritability will likely be due mostly to shared genetics but also partly to shared environment (as required by his second law: "The effect on the phenotype of being raised in the same family is greater than the effect of genes"; X[illegible]X). And if we

*To contact the authors, please write to: Carl F. Craver, Department of Philosophy, Washington University in St. Louis; e-mail: ccraver@wustl.edu. Mikhail Dozmorov, Department of Biostatistics, Virginia Commonwealth University; e-mail: mikhail.dozmorov@vcuhealth.org). Mark Reimers, Neuroscience and Bioengineering Program, Michigan State University; e-mail: reimersm@msu.edu. Kenneth S. Kendler, Virginia Institute for Psychiatric and Behavior Genetics and the Department of Psychiatry; e-mail: kenneth.kendler@vcuhealth.org.

†Thanks to Niklas Andersson, Lindley Darden, Susan Fitzpatrick, Mark Povich, and two anonymous referees for comments and to Frederick Eberhardt especially for discussing s-t cuts. This project is funded in part by the John Templeton Foundation via a Genetics and Human Agency Award. This work is solely the responsibility of its authors and does not necessarily represent the official view of the funder.

Philosophy of Science, 87 (December 2020) pp. 1–12. 0031-8248/2020/8705-00XX$10.00
Copyright 2020 by the Philosophy of Science Association. All rights reserved.

14869.proof.3d 1 Achorn International 10/31/20 06:45

gather a large sample of 25,000 roller coaster enthusiasts and compare their genomes to those of the roller coaster–phobes, we will surely find numerous SNPs (single nucleotide polymorphisms) enhanced for (i.e., significantly correlated with) liking and disliking roller coasters. The question is: Will the various SNPs this hypothetical study identifies tell us anything coherent about why some people love roller coasters and others do not?

This question is the heart of Turkheimer's criticism of behavioral and psychiatric genome-wide association studies (GWASs). At one level, Turkheimer notes that GWAS data thus far explain only a modest to moderate portion of the variance in behavioral and psychiatric GWAS (his third law).[1] But his deeper worry, what explains this disappointment, is that GWAS will tell us nothing intelligible or actionable about how genes and phenotypes are related; that is, it will not reveal the etiological mechanisms by which behavioral and psychiatric phenotypes arise and through which they might be enhanced, remedied, prevented, promoted, and modified.

Turkheimer predicts incoherence: "If the history of empirical psychology has taught us anything it is that correlations between causally distant variables cannot be counted on to lead to coherent explanatory models" (1998). Likewise, he claims that "the 'statistical models' of gene-trait associations lack any meaningful link to etiological models of the transmission of complex traits" (Turkheimer 2011, 232). Twenty years after he offered this dire prognosis, he delivered a sobering assessment: "the new era of human genomics has been a disappointment so far. . . . Human genomics can fill libraries with 'results' that fail to cohere in even partial explanations in terms of the 'biological roots' of psychiatric disorders and behavioral traits" (Turkheimer 2012, 44). Incoherence, he argues, is the "gloomy prospect" hanging over the future of behavioral and psychiatric GWAS.

Our primary aim is to clarify the coherence problem. In fact, Turkheimer's objection elides two distinct but interacting sources of incoherence requiring different solutions: *heterogeneity* and *complexity*. We examine each, describing the problem, explaining why it persists, and considering strategies for addressing it.

2. Heterogeneity and Coherence. One source of incoherence in GWAS is heterogeneity: there is likely individual variability in the behavioral/psychiatric phenomenon itself (phenotypic heterogeneity), in its underlying mechanism (mechanistic heterogeneity), and in the genetic differences implicated in its etiology (genetic heterogeneity). People avoid roller coasters for different reasons: some are acrophobic, claustrophobic, or emetophobic; others avoid crowds or hate noise; others are prone to motion sickness. Among those who avoid coasters because of motion sickness, we again confront

1. See, e.g., Maher (2008) and responses by Bourrat and Lu (2017).

diversity: this might be due to differences in the vestibular system, in the visual system, sensory integration mechanisms, the gut, and perhaps even in the gut biome. And supposing that we have identified one such mechanism (e.g., vestibular effects), we again confront many potential pathways from gene to phenotype: malformation of the semicircular canals, the neurons in the vestibular ganglion, or the conformation of mechanosensitive channels. At every level and at each stage in the etiological mechanism, we find reasons to expect multiple realization or *equifinality*: there are combinatorially many, many ways one might dislike roller coasters. Surely this diversity will be represented in our large sample of roller coaster avoiders. Perhaps psychiatric disorders are heterogeneous like that too.

More abstractly: GWASs are premised on the ability to extract information about etiology from aggregate data about populations. Saying a SNP is *associated* tells us nothing about any particular person's genome; it measures a SNP's frequency in a population of individuals. In current generation GWAS, populations are defined only relative to a phenotypic characterization that glosses over plausible phenotypic, mechanistic, and genetic variation. This seems especially likely in behavioral and psychiatric domains, where taxonomies are notoriously fluid, controversial, and user dependent. Such taxonomic problems bedevil the behavioral sciences generally.

3. Heterogeneity as a Consequence of Development. A more troubling possibility, though, is that this fluidity, controversy, and inconsistency is the consequence of biological variability rather than, or in addition to, scientific immaturity. Organisms develop through "soft-construction" processes (Clark 1998); they are tinkered and cobbled together through decentralized mechanisms that are selective rather than instructive. They are shaped by end point feedback and the reduction of error signals, and the end point might be reached in many ways, as the organism makes flexible use of available resources at the required moments. Soft construction allows for a population of systems that *satisfice* (i.e., do tolerably well most of the time) but do so with very different mechanisms.

Eve Marder's work on the crustacean stomatogastric ganglion provides a glimpse into this soft-construction process and its implications for thinking about how heterogeneity might contribute to the incoherence of GWAS (see Prinz, Bucher, and Marder 2004; Marder, Goeritz, and Otopalik 2015; O'Leary and Marder 2016; Nassim 2018). The stomatogastric ganglion contains a pattern generator that drives the peristaltic wave carrying food through the intestine. This pattern generator comprises 14 neurons with relatively stereotyped inhibitory connections to one another. Collectively, they produce a regular bursting pattern that drives the intestine's musculature.

For years, Marder assumed that the ion channels involved in the pyloric rhythm must be finely tuned to produce the burst pattern. After repeated

frustration, she began to question this assumption. She and her colleagues built about 20 million simulation models of the pattern generator, including detailed settings for individual ionic conductances. She found that about a half million, some with tunings very different from the others, produced functional rhythms. When Prinz and her colleagues (2004) returned to the electrophysiology, armed with this space of possible conductance tunings, they indeed found examples of the functional variants their simulation models had identified.

The implications of this work for behavioral and psychiatric GWAS have not previously been made explicit. First, the stomatogastric ganglion of the crustacean is a primitive system (by any measure) with a stereotyped function; one might expect hardwired construction here if anywhere. If we find soft construction here (and we do), we should surely expect it in more "complex" mechanisms and systems, especially those, such as the neocortex, selected for learning and plasticity. The cognitive structures thought to be involved in the etiologies of psychiatric disorders (e.g., executive control, impulsivity, hallucinations, emotional regulation) are soft constructed and so likely massively heterogeneous. This heterogeneity is then reflected back in the GWAS.

The second implication becomes apparent when we consider how crabs with phenotypically type-identical pyloric rhythms break down when placed in conditions of stress, such as rising temperatures. The soft-construction process "searches" for a way of tuning conductances that works well enough in a large enough range of conditions. O'Leary and Marder (2016) found homeostatic mechanisms that adjust conductances to maintain robust burst patterns with changes in temperature, at least up to a point. Where that point lies, however, and what happens beyond that point varies, often considerably, one crustacean to the next. Phenotypically identical patterns are underwritten by heterogeneous mechanisms that fail differently in different individuals.

Perhaps psychiatric disorders are similar. If so, we should expect considerable individual variation in the way that soft-constructed systems fail (perhaps explaining some of the diagnostic complexity in psychiatry). And we should expect different individuals to be more or less robust in surviving life stressors without a major psychiatric episode (see, e.g., Fergus and Zimmerman 2005). But the underlying mechanistic story delivers the primary implication for behavioral and psychiatric GWAS: these different resiliencies might be grounded in different ways of reaching a homeostatic set point in the regulation of ion channels across a network, something the whole system has settled into as a satisficing solution to the challenges of development. The space of satisficing solutions compatible with daily function might be very large, leading to a population of individuals each of whom has a different susceptibility to malfunctioning.

This simple experimental system shows how soft construction leads to heterogeneous populations that are differentially susceptible to breakdown in the face of different environmental challenges. If genes are related to behavioral and psychiatric traits via soft-construction mechanisms, like the stomatogastric ganglion of the crustacean, GWAS will necessarily confront a massive heterogeneity problem.

4. Addressing Heterogeneity. What might more tolerable heterogeneity look like? Consider a simple etiological schema with four variables arranged in a causal sequence, in which each antecedent is sufficient for the next:

$$A \rightarrow B \rightarrow C \rightarrow D.$$

A maximally homogeneous population would have a globally overlapping pattern: each individual instantiates the same sequence. Cases of the flu are like this, up to a point. But perhaps AIDS is the better analog: individuals with AIDS all have etiological mechanisms that overlap at key stages of the mechanism (e.g., infection of T cells) but that diverge from that point on. In HIV infection, there is only *local pattern overlap* among individuals with the disorder. Individuals share only some of the components in the etiology or the constitutive mechanisms and diverge elsewhere.

When overlap fails, the appropriate scientific strategy is to *split* the putative phenotypic kind. The sample population is decomposed into subpopulations representing distinct etiological types with varying degrees of local or global overlap. As discussed below, such subpopulations would show up as modules in the overall network structure of the developmental system.

Solving the heterogeneity problem requires finding areas of local overlap that are sufficiently common among independent subtypes that they apply broadly across the variants. Given that soft construction likely produces etiological heterogeneity, where might one look for useful overlap? One might investigate the developmental mechanisms themselves; perhaps these offer common targets across diverse mechanisms. Or perhaps the homeostatic mechanisms designed to keep the satisficing solution in place offer local points of similarity or overlap. Perhaps that mechanism can be tuned, setting the thermostat to a different set point and allowing the system to settle into a new dynamic. Perhaps we should think about rendering these mechanisms plastic and allowing them to settle into a different "tuning." Perhaps the power of intelligibility and control lies not in the particular configuration of the system but in the causal mechanisms that produce, underlie, and maintain those systems over time.

The problem of heterogeneity is felt most acutely in the discovery process and in the search for useful targets of intervention. Concerning discovery, a finding that over 100 SNPs are statistically associated with the disorder (such as the findings from the GWAS consortium working groups on schizophrenia

[SWGPGC 2014] and major depressive disorder [Wray et al. 2018]) likely indicates a phenotypically similar but mechanistically diverse population (like the varieties of roller coaster phobia discussed above). Perhaps different enhanced SNPs reflect different subpopulations, and perhaps (as explored below) they reflect different semi-independent modules in a larger network. Concerning control, an effective intervention against one variant might not work against others. We consider below how the search for causal bottlenecks might help address the heterogeneity problem. Yet the enterprise as a whole faces a challenge: as genetic and developmental heterogeneity increases, it might become more cost effective to seek local overlap in environmental mechanisms and individual-level interventions than in the genetic factors.

5. Complexity and Coherence. A second source of incoherence in GWAS is the causal complexity of the mechanisms: the variables and interactions are so numerous, diverse, and reticulate that one might despair of finding any identifiable target for intervention or any identifiable interface in terms of which to render the community further intelligible. Rational design for control over a system might proceed by learning how to intervene on a manageable set of variables that together provide sufficient leverage over the mechanism's behavior to alter it in the direction of the goal. Our very idea of "intelligibility" often turns on identifying in something difficult to understand a set of reliable interfaces, separating regular components, and decomposing a complex phenomenon into less complex, independently intelligible but interacting subcomponents (Haugeland 1998). Control and understanding are deeply connected in the scientific mind (Woodward 2003).

Do not ask: Will GWAS give us a complete explanation for psychiatric disorders? Genes are never sufficient for phenotypes, and no scientific method ever does it all. Ask instead: How might GWAS contribute to the search for intelligible and controllable mechanisms? What are the barriers to that goal? And how might they be crossed?

Thus framed, it is clear that a central constraint on the evidential power of GWAS is that it focuses on one kind of cause (genes, broadly construed), with no direct information about the cellular, physiological, organismic, and environmental causes with and through which genes act. The power of GWAS lies in its use of large sample sizes and high throughput genotyping methods. Because it is labor and time intensive to collect standardized data on behavioral and environmental cofactors causally relevant to psychiatric phenotypes, current generation GWASs typically neglect and, in some cases, crowd out projects designed to study these variables (Tabery 2018). GWAS gives us one kind of piece in the puzzle box, perhaps some corner pieces, but we are left to infer how the rest of the puzzle fits together around them, and the more complex a mechanism is, the more perilous this inference becomes.

A possible strategy for addressing this limitation, which we explore in the next two sections, is to forward chain inferentially (in the sense of Craver and Darden 2013) from enhanced SNPs to genes and from genes to molecular mechanisms using databases of gene-protein relationships and protein-protein interactions (PPIs; see, e.g., Reimers et al. 2019). This sort of forward chaining has, in fact, delivered important outcomes elsewhere, perhaps most notably the GWAS of age-related macular degeneration, a leading cause of blindness and vision impairment in adults (see Fritsche et al. 2016). A 2005 GWAS identified 19 genes as genome-wide significant for age-related macular degeneration. Four of these, including three of the most significant hits, were from a previously unsuspected complement pathway, suggesting a possible target for intervention. Therapies targeting many parts of the complement pathway are now in clinical trials. This is an exemplar of the sort of forward-chaining insight one might hope GWAS will deliver about downstream mechanisms and therapeutic targets, even if it falls short of telling us "everything" we want to know. GWAS might, in fact, point to some unsuspected reasons people do not like roller coasters.

Any such effort at forward chaining from current GWAS data must cross the hurdle of inferring implicated genes from enhanced SNPs. GWAS uncovers the sequence in the genome by sampling key "tag SNPs" and using linkage disequilibrium (genome regions with nonrandom correlations among SNPs) to infer potential causal SNPs. Enhancement of a SNP in GWAS is an indicator that some SNP in linkage disequilibrium with this tag SNP is correlated with (indeed causally relevant to) the disorder. Importantly, SNPs are not genes and may not even be parts of genes and may not even be parts of gene regulatory units. So, the inferential step from SNP to gene requires picking out the relevant genetic differences from the sea of everything in linkage disequilibrium with that SNP. From there, it is an additional challenge to figure how this all makes a difference to proteins and their interactions.

Having thus associated tag SNPs with genes, one might then associate the implicated genes with proteins and, appealing to databases of known protein-protein connections (see Bajpai et al. 2019), forward chain the likely networks of connectivity among the gene products. Measures such as the minimum s-t cut, discussed below, are directly sensitive to whether one's variable set is complete. We know the databases are incomplete and often do not known how incomplete they are or in what respects (see, e.g., Menche 2015). Yet, on the assumption that the databases have identified many of the key players, one can begin to formulate meaningful network estimates of some important factors. One can, for example, ask whether the items in the set of identified genes and products interact with one another substantially more than they do with the proteins associated with random SNPs not identified in the GWAS. This could be seen as a network measure for the extent to which the SNPs and their products form a "community" (Lancichinetti,

Radicchi, and Ramasco 2010; Fortunato and Hric 2016). Kendler (2013, 1062, fig. 2a) visualizes this idea as a spread of SNPs and copy variants with edges joining most of them into a connected graph.

Two limitations of this approach stand out. One is that these measures are blind to *explanatory relevance*, that is, to the intelligible link between enhanced SNPs and phenotypes, the proper guide to something exploitable for control (Woodward 2003). The majority of PPI databases contain all the interactions in which these proteins engage, not only the ones that are relevant to the tissue- or organ-specific phenotype. The detected community structure might be due entirely or primarily to irrelevant interactions. We have only recently started to scratch the surface of tissue-specific gene regulation (Kitsak et al. 2016), but until we have a better understanding of relevant interactions (perhaps brain-specific interactions, in the case of psychiatric disorders), our understanding of community structures will be partial and potentially misleading.

Second, a set of variables might form a community through a dense network of interactions that cannot be decomposed into intelligible subunits interacting through well-defined interfaces. The connections among community members might be thick and dense, lacking any localized bottleneck at which one might reasonably hope to target one's therapeutic interventions. Communities can be causal thickets, with bewilderingly diverse and reticulate connectivity among components. If SNPs are connected to phenotypes only through dense causal thickets, the prospects for intelligibility and manageable intervention are gloomy.

That said, such ways of thinking about global community structure are potentially very instructive about downstream mechanisms inferred from GWAS findings. Next-generation tools for dealing with complexity ~~by searching for substructure in these communities~~ have begun to reveal modular substructure (Oti and Brunner 2007; Choobdar et al. 2019). Recent research reveals that networks formed by disease-associated genes are composed of highly interacting (coherent) subnetworks (Ghiassian, Menche, and Barabási 2015). Notably, the most coherent subnetworks were found for psychiatric and immune-related disorders, blood cholesterol, and anthropometric traits, for which high-powered GWASs are available. These subnetworks are strongly enriched in various pathways and gene ontologies (Choobdar et al. 2019), suggesting complementary molecular mechanisms associated with the same phenotype that may be therapeutically targeted (Lagunes-García et al. 2019). This provides some grounds for optimism, although characterization of disease-associated networks is still in its infancy.

The identification of community substructure (i.e., modularity within the overall community) might offer important insight into how to deal with the heterogeneity problem discussed above. A modest failure of overall community connectivity (fig. 2d in Kendler 2013, 1062) might still point to coherent

networks in subpopulations (as in Kendler's figs. 2b and 2c). Alternatively, such modular substructure might be taken to suggest that the mechanism requires "multiple hits" to different more or less isolated mechanisms. In short, details about community structure, arrived at by forward chaining from SNPs, might well suggest novel hypotheses about etiological mechanisms. We next consider a way of understanding the sort of "causal bottlenecks" or "interfaces" that would constitute one kind of solution to the problem of complexity.

6. Addressing Complexity with Bottlenecks. The complexity challenge to coherence is that GWAS is unlikely to reveal intelligible causal structures, that is, identifiable interfaces in the mechanism. Overall community structure, we argued in section 5, cannot address that challenge. What is desired, both for understanding the system and for defining appropriate intervention targets, is a set of well-defined, high-traffic interfaces in the mechanisms that produce, underlie, or maintain the trait in question. We call these "bottlenecks" in the productive flow of the mechanism.[2]

Intuitively, a bottleneck is an edge, or set of edges, on which many independent causal paths converge. Such bottlenecks are promising targets for intervention and concentrated sources of information about downstream effects because much of the mechanism's complexity funnels through this common channel. For example, a complex supply chain converges on the auto plant; control over factories is control over car production. One way to address Turkheimer's criticism would be to trace directed causal paths from their beginnings (their "sources" in genes and environment) to their ends (the "sink" in phenotype) and then to identify places where many independent mechanisms course through the narrowest channels.

This intuitive idea connects with a concept from network science: a *minimum s-t cut*. If we represent the causally complete etiological mechanism here as an arrangement of causal paths between sources (s: genes and environments) and sinks (t: phenotypic traits), what we want in the ideal chokepoint is that edge or set of edges through which all the causal "flow" from beginning to end must pass. The ideal chokepoint in a causal graph is a minimum cut: the fewest edges you have to cut to divide the graph in two, with one graph at the beginning and one at the end. An advantage to conceptualizing chokepoints in this way is that there are tractable algorithms for finding the minimum cut in directed and weighted networks (such as the Ford-Fulkerson algorithm; Cormen et al. 2001).

2. Steele (1999) emphasizes the importance of such points of convergence in making inferences from model organisms to targets. Jones (2014) discusses the same idea under the heading of "bowtie" structures in causal networks; Tieri et al. (2010), e.g., apply a bow tie analysis to the immune system.

A chokepoint clearly need not be a member of the minimum cut to be useful. The key idea is graded: bottlenecks are nodes and edges that represent crucial control points in mechanisms on which many independent causal influences converge and from which a wide range of causal influences follows. In fixing a variable, V, to a given value v, one renders all the ancestors of V irrelevant to the occurrence of the effect in question, except insofar as those variables act via a pathway that does not pass through V. The probability of the effect given V and its ancestors is the same as the probability of the effect given V alone (V *screens off* its ancestors). And as a common cause, V affords both control and prediction of descendent variables in the directed graph. Perhaps the edges connecting modules within a network afford a more local form of bottleneck that could be useful even if it falls short of the minimum s-t cut (a local minimum perhaps).

The structure of the network, and the location of bottlenecks, depends fundamentally on variable choice and, in particular, how coarse- or fine-grained the variables are. A bottleneck might include a downstream protein, such as a mechanoreceptor in the ear, with different SNPs jointly relevant to the final common pathway of amino acid sequence or gene regulation. More abstractly, perhaps many enhanced SNPs contribute to different proteins that all participate in a common higher-level mechanism, such as transduction in the semicircular canals. More abstractly still, motion sickness might be a higher-level node common across distinct lower-level etiologies for roller coaster–phobia involving the gut, the inner ear, and the eyes. Indeed, many of the SNPs identified in Hromatka et al.'s (2015) GWAS of motion sickness appear to point to balance, cranial development, and glucose metabolism systems. Perhaps control over these nodes offers leverage over the phenotype, a connection discoverable by chaining forward from GWAS. The search for coherence in complex causal networks, or modular organization in the global community of interactions, can proceed at many "levels of organization" or gains of analysis simultaneously, wherever one can find a chokepoint to exploit.

7. Conclusion. Our goal has been to reconstruct Turkheimer's coherence challenge with an eye to addressing it. We distinguish two potential sources of incoherence: heterogeneity and complexity. The problem of heterogeneity (sec. 3) arises not only from limits of current-generation science but also, more fundamentally, from the ubiquity of soft construction in development. This problem can be addressed by finding regions of local overlap (sec. 4) or by subdividing the population into distinct etiologies (sec. 5). The problem of complexity is being addressed using databases of PPIs to forward chain from enhanced SNPs out to likely mechanisms. Current generation approaches face obstacles of relevance and of identifying community structure (sec. 5 above). Such approaches do not yet address the complexity problem directly. To do so, one might search for bottlenecks in directed graphs (sec. 6

above). The idea of a minimum s-t cut offers a concise definition of a bottleneck: a place where diverse streams of causal current pass through narrow chasms in the causal structure of the world. These are important places to look in the search for intelligible explanations and useful targets of intervention. By recasting Turkheimer's challenge in these terms, we hope thereby to expose these challenges and their solutions to further clarification through collective philosophical criticism. In this Turkheimer is correct: whether GWAS can surmount those challenges will affect fundamentally whether it is perceived in retrospect as worth the considerable investment we have made in it.

REFERENCES

Bajpai, Akhilesh Kumar, Sravanthi Davuluri, Kriti Tiwary, Sithalechumi Narayanan, Sailaja Oguru, Kavyashree Basavaraju, Deena Dayalan, Kavitha Thirumurugan, and Kshitish K. Acharya. 2019. "How Helpful Are the Protein-Protein Interaction Databases and Which Ones?" bioRxiv. https://doi.org/10.1101/566372.

Bourrat, Pierrick, and Qiaoying Lu. 2017. "Dissolving the Missing Heritability Problem." *Philosophy of Science* 84 (5): 1055–67.

Choobdar, Sarvenaz, et al. 2019. "Open Community Challenge Reveals Molecular Network Modules with Key Roles in Diseases." bioRxiv. https://doi.org/10.1101/265553.

Clark, Andy. 1998. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. 2001. "The Ford-Fulkerson Method." In *Introduction to Algorithms*, 2nd ed., 651–64. Cambridge, MA: MIT Press.

Craver, Carl, and Lindley Darden. 2013. *In Search of Mechanisms: Discoveries across the Life Sciences*. Chicago: University of Chicago Press.

Fergus, Stevenson, and Marc A. Zimmerman. 2005. "Adolescent Resilience: A Framework for Understanding Healthy Development in the Face of Risk." *Annual Review Public Health* 26:399–419.

Fortunato, Santo, and Darko Hric. 2016. "Community Detection in Networks: A User Guide." *Physics Reports* 659:1–44.

Fritsche, Lars G., et al. 2016. "A Large Genome-Wide Association Study of Age-Related Macular Degeneration Highlights Contributions of Rare and Common Variants." *Nature Genetics* 48:134–43.

Ghiassian, Susan Dina, Jörg Menche, and Albert-László Barabási. 2015. "A DIseAse MOdule Detection (DIAMOnD) Algorithm Derived from a Systematic Analysis of Connectivity Patterns of Disease Proteins in the Human Interactome." *PLoS Computational Biology* 11: e1004120. https://doi.org/10.1371/journal.pcbi.1004120.

Haugeland, John. 1998. *Having Thought*. Cambridge, MA: Harvard University Press.

Hromatka, Bethann S., Joyce Y. Tung, Amy K. Kiefer, Chuong B. Do, David A. Hinds, and Nicholas Eriksson. 2015. "Genetic Variants Associated with Motion Sickness Point to Roles for Inner Ear Development, Neurological Processes and Glucose Homeostasis." *Human Molecular Genetics* 24:2700–2708.

Jones, Nicholaos. 2014. "Bowtie Structures, Pathway Diagrams, and Topological Explanation." *Erkenntnis* 79 (5): 1135–55.

Kendler, Kenneth S. 2013. "What Psychiatric Genetics Has Taught Us about the Nature of Psychiatric Illness and What Is Left to Learn." *Molecular Psychiatry* 18:1058–66.

Kitsak, Maksim, Amitabh Sharma, Jörg Menche, Emre Guney, Susan Dina Ghiassian, Joseph Loscalzo, and Albert-László Barabási. 2016. "Tissue Specificity of Human Disease Module." *Scientific Reports* 6:35241. https://doi.org/10.1038/srep35241.

Lagunes-García, Gerardo, Alejandro Rodríguez-González, Lucía Prieto-Santamaría, Eduardo P. García del Valle, Massimiliano Zanin, and Ernestina Menasalvas-Ruiz. 2019. "DISNET: A Framework for Extracting Phenotypic Disease Information from Public Sources." bioRxiv. https://doi.org/10.1101/428201.

Lancichinetti, Andrea, Filippo Radicchi, and José J. Ramasco. 2010. "Statistical Significance of Communities in Networks." *Physical Review* E 81:046110. https://doi.org/10.1103/PhysRevE.81.046110.

Maher, B. 2008. "The Case of the Missing Heritability." *Nature* 456:18–21.

Marder, Eve, Marie L. Goeritz, and Adriane G. Otopalik. 2015. "Robust Circuit Rhythms in Small Circuits Arise from Variable Circuit Components and Mechanisms." *Current Opinion in Neurobiology* 31:156–63.

Nassim, Charlotte. 2018. *Lessons from the Lobster: Eve Marder's Work in Neuroscience*. Cambridge, MA: MIT Press.

O'Leary, Timothy, and Eve Marder. 2016. "Temperature-Robust Neural Function from Activity-Dependent Ion Channel Regulation." *Current Biology* 26:2935–41.

Oti, Martin, and Han G. Brunner. 2007. "The Modular Nature of Genetic Diseases." *Clinical Genetics* 71:1–11.

Prinz, Astrid A., Dirk Bucher, and Eve Marder. 2004. "Similar Network Activity from Disparate Circuit Parameters." *Nature Neuroscience* 7:1345–52.

Reimers, Mark A., Carl Craver, Mikhail Dozmorov, Silviu-Alin Bacanu, and Kenneth S. Kendler. 2019. "The Coherence Problem: Finding Meaning in GWAS Complexity." *Behavior Genetics* 49:187–95.

Simon, Herbert A. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106:467–82. q9

Steel, Daniel. 2008. *Across the Boundaries: Extrapolation in Biological and Social Sciences*. Oxford: Oxford University Press. q10

SWGPGC (Schizophrenia Working Group of the Psychiatric Genomics Consortium). 2014. "Biological Insights from 108 Schizophrenia-Associated Genetic Loci." *Nature* 511:421–27.

Tabery, James. 2018. "All of Us Are Getting Scammed: A Bait-and-Switch at the Heart of the Most Expensive Genetics Project in Decades." Paper presented at PSA 2018, November 3.

Tieri, Paolo, Andrea Grignolio, Alexey Zaikin, Michele Mishto, Daniel Remondini, Gastone C. Castellani, and Claudio Franceschi. 2010. "Network, Degeneracy and Bow Tie: Integrating Paradigms and Architectures to Grasp the Complexity of the Immune System." *Theoretical Biology and Medical Modelling* 7 (August): 32. https://doi.org/10.1186/1742-4682-7-32.

Turkheimer, Eric. 1998. "Heritability and Biological Explanation." *Psychological Review* 105:782–91.

———. 2000. "Three Laws of Behavior Genetics and What They Mean." *Current Directions in Psychological Science* 9:160–64.

———. 2011. "Still Missing." *Research in Human Development* 8:227–41.

———. 2012. "Genome Wide Association Studies of Behavior Are Social Science." In *Philosophy of Behavioral Biology*, ed. Kathryn S. Plaisance and Thomas A. C. Reydon, 43–64. Dordrecht: Springer.

Woodward, James. 2003. *Making Things Happen*. Oxford: Oxford University Press.

Wray, Naomi R., et al. 2018. "Genome-Wide Association Analyses Identify 44 Risk Variants and Refine the Genetic Architecture of Major Depression." *Nature Genetics* 50:668–81.

QUERIES TO THE AUTHOR

Q1. Au: Your article has been edited for grammar, clarity, consistency, and conformity to journal style. Please read the article to make sure that your meaning has been retained. Note that we may be unable to make revisions that conflict with journal style or create grammatical problems. Thank you.
Q2. Au: Please provide page numbers for Turkheimer's quoted first and second laws to replace the two XXX placeholders here.
Q3. Au: Please provide page number for quote ("If the history of . . .").
Q4. Au: Please revise introductory clause here ("Twenty years after he offered this dire prognosis"), which follows a quote from 2011 with one from 2012 (i.e., 1 year later, not 20). NB: This is also not 20 years after the first quotation in the paragraph from 1998.
Q5. Au: Menche 2015 not in reference list; please add, or omit citation here.
Q6. Au: Replaced parenthetical "a" with "fig. 2a" here, assuming you're referring to Kendler's figure on page 1062, OK? If not, please clarify meaning.
Q7. Au: As above (see query 6), replaced parenthetical letters with figure panels here, assuming you're referring to Kendler's figure on page 1062, OK? If not, please clarify meaning.
Q8. Au: Footnote 2: Steele 1998 not in reference list; please add, or omit citation here.
Q9. Au: Simon 1962 not cited in text; please cite, or omit from reference list.
Q10. Au: Steel 2008 not cited in text; please cite, or omit from reference list.