An Open Publishing Platform of
the Organization for Human Brain Mapping
APERTURE NEURO

ORIGINAL RESEARCH ARTICLE

Behavioral phenotypes, stochastic processes, entropy, evolution, and individual variability: Toward a unified field theory for neurodevelopment and psychopathology

Tonya White, MD, PhD

Section on Social and Cognitive Developmental Neuroscience, National Institutes of Mental Health, Bethesda, MD, USA

Twenty years ago during my postdoctoral research fellowship, I remember my mentor, Nancy Andreasen, telling me that when they first started neuroimaging studies in schizophrenia, their expectation was that they would "find the lesion," identifying the underlying neurobiology of schizophrenia. They started with images acquired using computer tomography and measured surfaces using an overhead projector and would later develop computer-based algorithms with data acquired from MRI. While they did find group differences between patients with schizophrenia and controls (1), bringing results down to the individual has been elusive. So, here, we are as neuroimagers, some 40 years after Nancy's and other researchers initial pioneering studies assessing individuals with psychiatric disorders, reading the recently published article by Marek et al. (2) titled "Reproducible brain-wide association studies require thousands of individuals" and wondering what should we be doing differently? While there is a room to critique the Malek et al. study, for example, the lack of controlling for age, sex, and other demographic factors that would have enhanced the precision of the relationships and reduced bias, the primary issue remains. Whether we discuss the need for hundreds versus thousands of individuals is a moot point, as we have yet to bring results down to the level of the individual.

Large sample sizes are required for population-based epidemiological studies, where the goal is to measure the influence of potentially small effects at the level of the population (3,4). However, when the goal is to understand the relationships within the context of precision medicine, finding small effects with large sample sizes is moving us in the opposite direction. Since we know that cognition, emotion, and behaviors are inherent in the brain, how is it that we have had limited success in finding precise links between neuroimaging metrics and behavioral and cognitive phenotypes? It is not uncommon to read in the introduction of papers: "prior studies evaluating the underlying neurobiology have mixed findings." The paper by Marek et al. highlights that sample size plays a major role in these mixed findings in the literature. However, the important question is, given the severity of brain-based disorders such as autism, schizophrenia, and bipolar affective disorder, why do we not see consistent and replicable markers of these disorders in the brain? It is this question that dominates my own brain activity (or connectivity), and it is this question that I believe is the most important to discuss in this editorial. In doing so, I will discuss two competing themes, the clustering of symptoms consistent with psychopathology and, the second, the considerable heterogeneity within a disorder.

At the beginning of the 19th century, Emil Kraepelin understood that individuals with specific psychopathology tended to have symptoms that clustered together (5). What Kraepelin termed dementia preacox would later be defined as schizophrenia and was defined by similarities in symptom profiles. This clustering of symptoms would later be expanded in the Diagnostic and Statistical Manuals of Mental Disorders, with the fifth edition being the most recent (6). While all disorders have considerable heterogeneity, including differences in the type and severity of symptom domains and comorbid disorders, classic phenotypes of psychopathology do form clusters in the population. These core features are brain based and, thus we, as neuroscientists, should be able to identify their underlying neurobiology. Although the resolution of these features may be at a much finer scale than our existing neuroimaging modalities. It is also possible that different pathways exist that lead to similar behavioral phenotypes. If the latter, then adaptations in our image process algorithms or the use of machine or

deep-learning algorithms could potentially identify subgroups who share different pathways. In any case, larger sample sizes can potentially provide key information on a global scale regarding underlying mechanisms and pathways between brain, behavior, emotion, and cognition. Thus, one possibility for our lack of success for precision medicine using neuroimaging in psychiatry is heterogeneity. Specifically, similar symptoms between individuals and comorbid symptoms can both have different underlying brain mechanisms when using statistical approaches that assume similar pathways.

Given all the factors that take place during development, including genetic recombination, epigenetic processes, and environmental influences, variability between individuals should not be surprising, especially considering the evolutionary challenge of building and maintaining something as complex as the human brain. By last count, there were just over 21,000 protein-coding genes (7), of which 30% to 50% are expressed in the brain (8). Thus, these 7,000 to 10,000 genes drive all the different neurodevelopmental processes that result in the growth and differentiation of the variety of different brain tissues. These 7,000 to 10,000 genes encode processes that include cell division, neuronal migration, synaptogenesis, apoptosis, synaptic pruning and dendritic arborization, and the orchestration of events associated with the neurochemical development of the brain. This relatively small number of genes codes for the building and maintaining of the approximate 86 billion neurons and 50 trillion synapses. This is a tremendous engineering feat that is only possible with redundancy coupled with tightly controlled feedback loops (Figure 1).

Within the context of a developmental feedback loop, consider symmetric subdivision along the ventricular zone. The neuronal progenitor cells undergo symmetric division along the ventricular zone from approximately 2 to 6 weeks of fetal life; one cell becomes two, the two become four, four become eight, and so on. This exponential growth must require a tightly controlled mechanism to regulate the transition to the next developmental stage, since each additional cell division will result in a doubling in the number of neurons in the brain. The period of symmetric subdivision is followed by asymmetric subdivision and neuronal migration, which occurs between approximately 6 to 24 weeks of fetal life. Neuronal migration results in the formation of the six layers of the cortical sheath. Considering these two neurodevelopmental processes, there are not enough genes to encode the precise location and connectivity patterns of the billions of neurons and trillions of synapses in the brain. Thus, genetic factors initiate and end the developmental processes, likely also with specific boundary conditions during the process. However, within each process, there is inherent randomness. Brownian motion is inherently random and will influence each step in the process. Further, rates of diffusion are influenced by temperature, and thus environmental factors such as maternal infection could also alter neurodevelopmental processes. Given the large number of factors that can induce variability, it would seem that the most interesting and perplexing question is not why there is so much heterogeneity, but rather given the relatively small number of genes to control a highly complex process such as building and maintaining a brain, how is it that typical neurodevelopmental trajectories and symptom domains actually cluster together. Are there elements within the different start, control, and end points that show specific weaknesses that cause behaviors to cluster together? Do certain processes during development infer an evolutionary advantage, resulting in specific domains clustering together? And what role does chance play in this process?

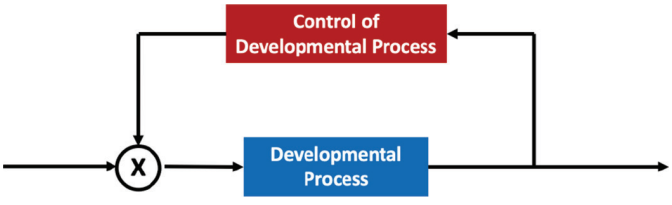

Fig. 1. Simple systems theory depiction of developmental feedback loop.

When animals are bred to have an identical genetic makeup and the environment is controlled to be as identical as possible, the level of individual variability continues to play a pronounced role in individual variability, which Klaus Gärtner labeled a "third component influencing development" (9). In fact, is diversity (heterogeneity) was so crucial from an evolutionary perspective that a genetic "random number generator" exists to ensure developmental diversity? Linneweber et al. (10) demonstrated clear stochastic processes in the dorsal cluster neurons of the visual system in drosophila and in humans the detailed patterns of the iris are completely random, even in twins. Thus, there exist enough independent mechanisms in development (i.e., cell division, neuronal migration, and synaptogenesis) in which the developmental entropy is adequately controlled, but within those borders of control, there are windows for diversity (11). Thus, we should not be at all surprised by the extent of individual variability, as it is inherently embedded in the processes of neurodevelopment and provides a strong evolutionary advantage, especially during times of adversity.

To summarize, there are consistent clusters of symptoms that are associated with psychopathology. These symptoms are brain based and thus we "should" be able to identify their underlying mechanisms. However, brains are unique to each individual, based on genetic, epigenetic, environmental, and random factors. Returning to the paper by Marek et al., I believe that increasing sample sizes are important for finding replicable patterns associated with specific behavioral phenotypes, for parsing

heterogeneity and that our field should move, similar to the field of genetics, to require replication studies when possible (12). Further, we should think outside the box for approaches different than BWAS to determine links between the brain and behavioral phenotypes. Machine and deep-learning algorithms may be beneficial in finding patterns that we mere mortals can miss. Another element not discussed, but that can be gleaned from the Marek et al. study, is that neuroimaging studies can provide important information regarding the level of individual variability for specific behavioral phenotypes. Some developmental processes may have tighter control of entropy than others. Individual variability is both a blessing and a curse. The blessing comes in the form of phylogenic and otologic diversity, which includes our evolutionary development into a species with emotions, cognitions, and behaviors. But there is also the curse of less optimal genetic combinations, rare variants, environmental toxins, or just bad luck that can impair function.

There is a beautiful balance in the creation of the human brain, driven primarily by genetic processes that both create the infrastructure to build different elements of the brain and to control the optimal timing to stop the individual building processes and controlling environmental and importantly, stochastic processes, or entropy gone wild. The balance must continue as well to ensure the propensity for further evolution, which also could be considered entropic time gone wild, which, considering climate change and the war in Ukraine, there is pause to wonder whether we on earth are not just in a local maximum but rather helping to foster universal entropy (13).

REFERENCES

1. Andreasen NC. Linking mind and brain in the study of mental illnesses: a project for a scientific psychopathology. Science. 1997 Mar 14;275(5306):1586–93.
2. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, et al. Reproducible brain-wide association studies require thousands of individuals. Nature. 2022 Mar;603(7902):654–60.
3. White T, Muetzel RL, El Marroun H, Blanken LME, Jansen P, Bolhuis K, et al. Paediatric population neuroimaging and the Generation R Study: the second wave. Eur J Epidemiol. 2018;33(1):99–125.
4. White T, El Marroun H, Nijs I, Schmidt M, van der Lugt A, Wielopolki PA, et al. Pediatric population-based neuroimaging and the Generation R Study: the intersection of developmental neuroscience and epidemiology. Eur J Epidemiol. 2013 Jan;28(1):99–111.
5. Géraud M. [Emil Kraepelin: a pioneer of modern psychiatry. On the occasion of the hundred and fiftieth anniversary of his birth]. L'Encephale. 2007 Sep;33(4 Pt 1):561–7.
6. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed.) [Internet]. American Psychiatric Association; 2013. Available from: https://doi.org/10.1176/appi.books.9780890425596
7. Willyard C. New human gene tally reignites debate. Nature. 2018 Jun 19;558(7710):354–5.
8. Naumova OY, Lee M, Yu S, Vlasova NV, Grigorenko EL. Gene expression in the human brain: the current state of the study of specificity and spatiotemporal dynamics. Child Dev. 2012;84(1):76–88.
9. Gärtner K. A third component causing random variability beside environment and genotype. A reason for the limited success of a 30 year long effort to standardize laboratory animals? Lab Anim. 1990 Jan;24(1):71–7.
10. Linneweber GA, Andriatsilavo M, Dutta SB, Bengochea M, Hellbruegge L, Liu G, et al. A neurodevelopmental origin of behavioral individuality in the Drosophila visual system. Science. 2020 Mar 6;367(6482):1112–9.
11. White TJH. Brain development and stochastic processes during prenatal and early life: you can't lose it if you've never had it; but it's better to have it and lose it, than never to have had it at all. J Am Acad Child Adolesc Psychiatry. 2019 Nov 1;58(11):1042–50.
12. White TJH, Frazier JA, Althoff RR, Novins DK. What the Journal of the American Academy of Child and Adolescent Psychiatry is looking for in neuroimaging submissions. J Am Acad Child Adolesc Psychiatry. 2021 Mar 1;60(3):324–8.
13. England JL. Statistical physics of self-replication. J Chem Phys. 2013 Sep 28;139(12):121923.

