Received: 19 April 2021 | Accepted: 20 April 2021

DOI: 10.1002/aur.2524

REPLY

Progress in autism research requires several recognition-definition-investigation cycles

Laurent Mottron[1,2]

[1]Department of Psychiatry and Addictology, University of Montreal, Montreal, Quebec, Canada

[2]CIUSSS-NIM Research Center, Rivière-des-Prairies Mental Health Hospital, 7070 Bvd Perras, Montreal, Quebec, Canada

Correspondence
Laurent Mottron, Riviere-des-Prairies Mental Health Hospital, 7070 blvd Perras, Montreal, QC, Canada H1E 1A4.
Email: laurent.mottron@gmail.com

Funding information
University of Montreal's M. & R. Gosselin Research Chair on Cognitive Neuroscience of Autism

We thank Drs. Constantino, Gillberg, and Lombardo for their comments on our proposal to initiate a "radical change" in the selection of subjects enrolled in research cohorts based on a reinterpretation of the heterogeneity of the current ASD category. We will first summarize our position, allowing us to clarify the points that gave rise to a number of erroneous interpretations. We will then address the following relevant questions: What can genetic epidemiology discover about autism without the help of clinical disciplines? Does the relationship between effect size and sample size exclude the possibility of heterogeneity diminishing effect sizes? How can we achieve an iterative heuristic cycle between the definition of autism and fundamental research that breaks the current deadlock?

REMOVING SOME AMBIGUITIES FROM THE "BACK TO PROTOTYPE" PROPOSITION

Prototype theory, applied to autism, describes the bottom-up emergence of a pattern of clinical signs in the brain of a person relative to a collection of individuals with some level of resemblance. This prototype grades the exemplars of clinical presentations according to their level of resemblance to each other and their greater or lesser distinction vis-à-vis other conditions. It allows clinicians to recognize autism before and independently of "top-down" verification that the signs listed in the criteria or standardized instruments are embodied in the person considered. It may secondarily reveal its aggregation with non-clinical variables, but the first step is purely phenotypic.

The main idea of our proposal is therefore that autism targeted by the current criteria and instruments is not a "singular unitary entity," but that these criteria prevent separating intrinsic from artificial variability. The problem is not heterogeneity *per se*, but the wrong kind of heterogeneity. The ASD category presents a level and nature of heterogeneity that challenges unsupervised subgrouping. Definitional autism leaves us with only a choice between an unusable and trivial category and the absence of any category, ("all autistic people are different") that is equally so. If we dismantle the ASD category based on this proposal, we expect that the resulting elements will vary in size and level of intra-class homogeneity and that, indeed, some of them may be less heterogeneous than those of the current ASD spectrum.

Such prototypes(s) are not a priori *defined* in our proposition because we consider that definitions, in this case, are intrinsically less specific than prototype recognition: you recognize a face, you do not define it. Thus, post-recognition candidate properties of the prototype include socio-communicative negative signs present in a specific time window and perceptual positive (not "sensory") signs, including the domains of behavioral spontaneous orientation. We expect a prototype-based delineation to strongly modify the common assumptions on language, intelligence, and adaptation of autistic people, while severely narrowing the reported prevalence of autism: prototypicality may or may not overlap with severity (Dawson, 2009). Ongoing work in our group suggests that several variables of a biological nature and several values of the specifiers aggregate within this prototype: low VIQ/QIP ratio, early pattern detection, absence of deleterious deletions, head-size deviation, sex

ratio, "bayonet" language profile (typical start, regression or plateau, recovery), and non-social language learning.

Our proposal consists of creating a distinct trend of clinical and fundamental research on cohorts limited to subjects of maximum prototypicality defined by the expert clinicians who enroll them. It concerns "truncating" the distribution of such prototypicality for autism research, and studying it separately from other, less prototypical presentations. The practical details requested by C. Gillberg suggest, however, that our proposition is slightly misunderstood; while its standards are yet to be clarified, training in expertise *is* clinical exposure. Anyone who has been sufficiently exposed to an enriched population is likely to classify a sample of individuals as decreasing in prototypicality by quantifying the representativeness of the person to his mental image of autism, as suggested by the "frank autism" (de Marchena & Miller, 2017) study. As age is a major factor of variability and attenuation of differential features in autism, we may have to select the research populations according to their clinical picture at an age at which the signs are temporarily stable and at which their initial phenocopies have deviated from this presentation. It is expected that the easiest and most reliable prototype will be detected among preschool children with a speech delay. A second prototype, emerging at school age, is probably needed for people without delayed language onset, corresponding with the Asperger-type presentation.

The number of experts and the level of agreement among them and with existing standardized tools are, in our proposal, the purpose of secondary studies, not their starting point. We are not proposing an initial screening, but rather post-hoc mapping based on such prototypes on cohorts selected using another principle. We do not expect populations selected following a "prototype" principle to be identical to each other everywhere on the planet. Such a focus on reliability is, for us, exactly what has to be temporarily suspended. There will be time in the future for the various cohorts to be used for replication once substantial discoveries have been made. Moreover, the cohorts currently available are not equally superimposable due to the width of the ASD category. The current inter-judge reliability of the instruments reflects consensus, not truth: reliability is neither necessary, nor sufficient for discovery, and may even prevent scientific progress. Reducing the focus on instrument-based reliability also introduces noise that we predict to be inferior and distinct from that introduced by a criteria-based definition of autism. Currently, to think "out of the box" in autism research requires bypassing the monopoly that ADOS-ADI, SRS, and AQ hold on its worldwide delineation.

We must place ourselves in a situation in which we will detect one or more groups whose members are more similar to each other than to other members of the ASD category within the populations to which clinicians are exposed. It is therefore a matter of returning to Kanner's situation (and Asperger's for people without speech onset delay) to give the recognition-definition cycle a second spin, which will require going back to recognition rather than definition. Thus, we are suggesting a "semi-directed" heuristic pathway, as Lombardo suggests, but based on ordinary intelligence, not first artificial intelligence. It was undoubtedly necessary to attempt to standardize diagnostic procedures 30 years ago. It is now necessary to correct for the resulting impasses.

SUCCESS AND LIMITS OF GENETIC EPIDEMIOLOGY IN DISCOVERING THE "CAUSE" OF AUTISM

Genetic epidemiology and molecular genetics have made it possible to disentangle de novo and familial genetic contexts within the current ASD category, which represents immense progress. One subgroup, non-syndromic autism, mostly familial, is relatively homogenous at the phenotypic and cognitive level, aside from the variation introduced by developmental transformations, expertise (e.g., savant syndrome), and language history; the other subgroup is a miscellaneous mix of phenotypes presenting a diminished non-verbal IQ and an indefinite series of genetic insults, mostly de novo. However, negatively defining non-syndromic autism does not specify its boundaries with other psychiatric conditions, which also have a familial temperamental aspect - hence the multiple studies finding so called "autistic traits" throughout the DSM and prematurely inferring a common mechanistic basis on their expression. Moreover, genetic epidemiology and the tools promoted by this discipline appear to be incapable of differentiating between predisposition and condition. Quantitative tools, such as the SRS or AQ/EQ/SQ, devalue this distinction due to their "quantitative" nature, which preselects dimensions that linearly vary in their object.

Constantino's comparison between modeling reciprocal socialization and hypertension is doubly illuminating, not because of the interest of the variable measured by the SRS, but because of its limits and the biases it introduces. While blood pressure is a natural variable that is measurable on a continuous basis, the reciprocal socialization construct is a fuzzy, multifaceted, culturally variable, and normatively defined construct. The comparison of autism to hypertension demonstrates that autism is reduced by the SRS to only one continuous variable. Such a level of reduction is analogous to reducing an object to its price, or an animal to its weight. It defines a logical space in which certain things can be said that are true and obey laws, but it does not allow us to understand the set of properties of the object, some of which possibly have more consequences than their price or weight.

The absence of a one-to-one link between the predictors and the condition, admitted by Constantino in his

response to our proposal, indicates that there is an entire topic of scientific inquiry missing between the two. Pleiotropism is not, in our opinion, a justification for devaluing the predictor/condition distinction on the pretext that the condition studied shows certain variability. Pleiotropic variation between twins cannot justify the level of generality and the "quantitative" nature of the concepts used to describe autism, exemplified by the SRS and its low level of specificity. Constantino suggests that our proposition "misses dimensionality". His notion of the "continuous distribution of autistic traits" misses, in turn, the distinction between control variables (continuous) and state variable (discontinuous) in a dynamic system.

The notion of pleiotropism only works well unidirectionally (figure 1 in Mottron, 2021). If used against the grain, it prevents the detection of possible natural categories or excludes them in advance. Genetic epidemiology is unable to limit the "wide range of variation" that Constantino believes is the right level of study for autism. The expression "arbitrary or expert-defined phenocopies" used to describe the variation within ASD according to the DSM-5 clearly indicates that in his thinking, the variations observed are pre-categorized as phenocopies because autism is necessarily comprised of a set of phenotypes produced by causative families of determinants. It a priori excludes that there could be, within such a definition of ASD, subgroups defined by a relatively unitary mechanism, regardless of their genetic determinants.

We believe that the explanation of ASD-related variability should be borrowed from a discipline other than genetics using another method, namely the clinic. To claim, in advance, that "the genetics of ASD is polygenic and incremental" and to assume that variations in phenotype are "stochastic" abandons any effort to explain what produces similar phenotypes, or prematurely devaluates their study. In another area of study, one would be critical of a scientific stance claiming that it is not worth studying the biological mechanisms of speciation because "all this is carbon."

The gap between Constantino's position and ours is illustrated by a preferential focus of 2 of the 4 Aristotelian causes of autism, its *material* cause (the genetic material) and its *final* cause (its evolutionist aspects), which are different from the two other causes that we investigate, its *formal* cause (the mechanistic nature of the autistic process, once triggered, responsible for its atypicality) and its *efficient* cause (the event triggering the bifurcation between predisposition and condition). Inheritance, clearly presented by Constantino as the cause of autism, is its *material* cause (in short, its genetics) but does not provide information about its *formal* cause or *efficient* cause. Although it is possible to classify the animals of Africa by increasing weight, as it is possible to classify humans by increasing level of reciprocal socialization, such ordering will teach us almost nothing about their individual biology or morphology. It will wrongly suggest that an elephant runs twice as fast as a zebra because it weighs twice as much or that some people are twice as autistic as others because their SRS score is two times greater. We will not be able to further assimilate the properties of the different entities because they are measured by the same variable and thus conclude that the weight of a zebra is measured in "elephant kilos," or vice versa, that the weight of an elephant is measured in "zebra kilos." The linearity of social communication is not a property of the object, it is introduced by its measure, the SRS.

This is, however, what the concept of 'autistic traits" assumes with the use of the term "autistic" to designate a cumulative, hidden variable of socialization, allowing the assertion that "autistic traits are continuously distributed in the general population" as dogma. The expression "autistic traits" is a contradiction "*in adjecto*." To formulate it as "what is true" (or "breeds true," which in fact does not appear to us to correspond to the type of transmissibility in question here) shows a very high level of confidence. The "social" construct has only modest descriptive value for autism, as there are multiple aspects of socialization that are typical in autism: emotional empathy, attachment, face recognition, and understanding agentive role. Its explanatory value is even more doubtful, whether one thinks of the savant-syndrome, for example, or head-size deviations. Autism is not a reciprocal socio-communicative problem but a variation in the way humans hierarchize, group, and generally process information structure and domains.

STATISTICAL ISSUES: SAMPLE SIZE, EFFECT SIZE, AND HETEROGENEITY

Our proposal to rethink autism research builds upon results from recent studies (e.g., Arvidsson et al., 2018; Idring et al., 2015), one of which is a decrease in effect size observed in autism neurocognitive studies over the last decades (Rødgaard et al., 2019). Lombardo questions whether there were indeed larger case/control effect sizes in samples studied in the past, which were likely less heterogeneous. As Lombardo points out, there are correlations between publication year, effect size, and sample size. There are at least two ways that these variables can be related to each other: (a) small studies can have inflated effect sizes through publication and reporting biases (Lombardo et al., 2019) and the trend toward newer studies using larger samples may have contributed to a trend toward smaller effect sizes. However, the analyses in Rødgaard et al. (2019) considered and controlled for differences in sample size. The finding of a decrease in effect size thus suggests a temporal effect *beyond* what can be explained by effect size inflation in early, small studies. (b) Large samples may have been obtained by sacrificing specificity or prototypicality in the recruitment process, which leads to larger heterogeneity and, in turn,

smaller effect sizes. Lombardo's objection to a suggestion that large samples are problematic in themselves may miss our argument that the problem actually stems from a by-product of large samples. Larger samples are always better, all else being equal. However, if all else is not equal, that is, if inclusion criteria are broadened to enable larger samples, heterogeneity will likely increase.

Lombardo also appears to question whether heterogeneous samples will in general result in smaller case/control effect sizes than more prototypical and thus more homogenous samples (i.e., samples that resemble each other to a larger degree than is currently the case with the autism population). Standardized mean differences (smd) (which are among the effect sizes most often used in case/control studies, and of which Cohen's *d* is the most common) are generally calculated as:

$$\text{smd} = (\text{mean}_1 - \text{mean}_2)/\text{std},$$

where mean_1 and mean_2 are the group means and std is the standard deviation of the samples. This definition means that if the samples become less heterogeneous (smaller standard deviation), the effect size increases, even if the group means stay constant. If the difference in mean values also increases, this causes an even larger increase in effect size. Because prototypical samples likely represent the most divergent segments of the current autism population and show less variation, thus smaller standard deviations, than the current autism population, it follows from the definition of standardized mean differences that less heterogeneous (e.g., highly prototypical) samples will result in larger standardized mean differences.

The decline in effect sizes in this context may also stem from the statistical tools used. Linear models fail to control for both the internal (between autistic) and external (between autistic and non-autistic) effect of heterogeneity as the sample size increases. Other methods may better capture the possible non-linear dependence between an autism diagnosis and the characteristics of individuals. Ultimately, we predict a convergent relationship, such that the effect-size becomes more stable as the size of the sample increases, allowing determination of the optimal size of the sample and maximizing the dispersion between homogeneous groups.

HOW CAN WE ACHIEVE AN ITERATIVE HEURISTIC CYCLE BETWEEN THE DEFINITION OF AUTISM AND FUNDAMENTAL RESEARCH THAT BREAKS THE CURRENT DEADLOCK?

Lombardo goes on to question whether the idea of prototypicality represents a "correct" model of autism. We see prototypicality as a cyclic heuristic pathway, not as a model. We do not suggest a return to prototypicality because it is necessarily a means to sharply delineate a categorically distinct clinical group, but rather because it does not exclude it a priori. Relatively unitary threshold-type mechanistic neurodevelopmental mechanisms may account for our ability to recognize autism, both in itself and in contrast to other neurodevelopmental conditions. Above all else, it is a way to break the current stalemate in autism research, in which heterogeneity appears to have become self-ascertaining and self-amplifying (Figures 1 and 2).

Lombardo and Constantino question why prototypicality should be defined at the behavioral and cognitive level rather than at the neural or genetic level. Conceptualization and investigation of psychiatric conditions have historically been derived from an initial description based on behavioral and cognitive features due to the observability of patterns on these levels in a clinical setting. Although the co-segregation of biological features with clinically defined patterns is expected, the

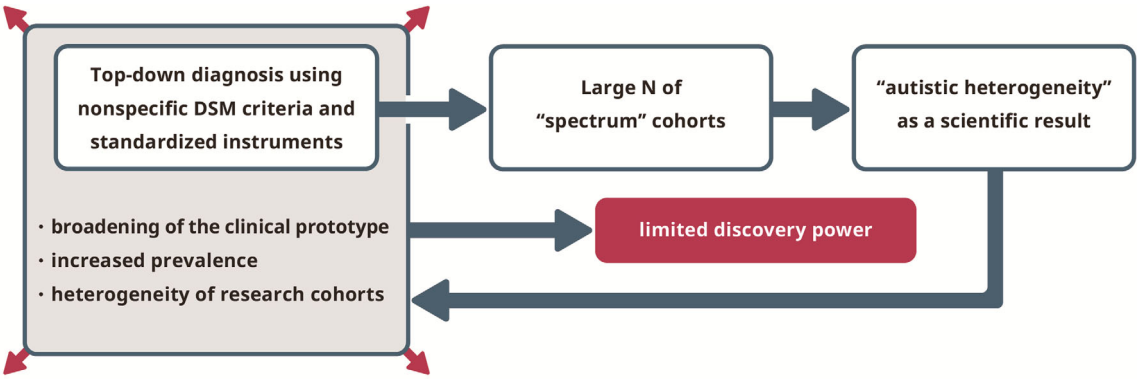

FIGURE 1 A model for how a self-amplifying cycle may result from the use of checklist criteria for diagnostic purposes and consecutive inclusion in autism research cohorts. The allowance of unequal levels of autism prototypicality compatible with an "autism spectrum disorder" diagnosis may produce results for which heterogeneity is attributed to the essence of the condition under study. This favors the constant diminution of the prototypicality required to obtain an autism diagnosis, an increase in prevalence, and a decrease in effect size in studies comparing autistic to non-autistic controls

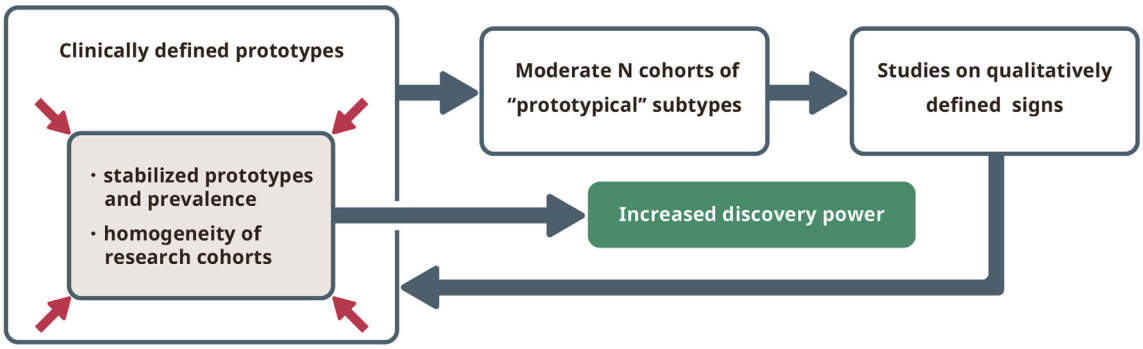

FIGURE 2 The hypothetical virtuous circle of phenotype refining. An initial clinical stratification results in a phenotypically homogeneous cohort, allowing comparison and specificity studies among prototype-based subgroups. The resulting characterization of autistic signs is re-entered into the phenotype description to increase its specificity and, ultimately, its discovery power

second is heuristically anterior to the first. Autism was first clinically recognized, then defined, and scientifically investigated. Research is currently plateauing (Figure 1), and we propose initiating another turn in the clinical-definition-neurobiological cycle (Figure 2) to foster progress in our understanding of this human variant.

ACKNOWLEDGMENTS

Although this article and the response to the comments it elicited benefited from discussions among researchers, clinicians, and students in the Montreal group, they may, however, not endorse its final content. We would particularly like to thank EM Rødgaard (U of Copenhagen), Michelle Dawson (U de Montreal), (J. Rabot (U. Strasbourg), Z. Abderrahim (U de Montreal), B. Bernhardt (U McGill), D. Bzdok (U McGill), G. Dumas (U de Montreal), C. Jacques (U de Montreal), R. Joober (U McGill), D. Gagnon (U of Montreal), S. Belleville (U of Montreal), I Dalahi (U Laval), F. Lesperance (U of Montreal), D. Ilies (U of Montreal), P. Grégoire (U of Montreal), EF Bui-Xuan (U of Montreal) S. Potvin (U de Montreal), and I. Soulieres (U de Montreal). The University of Montreal's M. & R. Gosselin Research Chair on Cognitive Neuroscience of Autism had no role in the study design, collection, analysis, or interpretation of the data, writing of the manuscript, or the decision to submit the paper for publication.

CONFLICT OF INTEREST

Laurent Mottron declares that he has no conflict of interest.

AUTHOR CONTRIBUTIONS

Laurent Mottron conducted literature searches, provided summaries of previous research studies, and wrote the final manuscript.

ORCID

Laurent Mottron https://orcid.org/0000-0001-5668-5422

REFERENCES

Arvidsson, O., Gillberg, C., Lichtenstein, P., & Lundström, S. (2018). Secular changes in the symptom level of clinically diagnosed autism. *Journal of Child Psychology and Psychiatry*, *59*(7), 744–751. https://doi.org/10.1111/jcpp.12864

Dawson, M. (2009). *Notes on autism severity and the DSM-V*. https://autismcrisis.blogspot.com/2009/06/notes-on-autism-severity-and-dsm-v.html.

de Marchena, A., & Miller, J. (2017). "Frank" presentations as a novel research construct and element of diagnostic decision-making in autism spectrum disorder. *Autism Research*, *10*(4), 653–662. https://doi.org/10.1002/aur.1706

Idring, S., Lundberg, M., Sturm, H., Dalman, C., Gumpert, C., Rai, D., Lee, B. K., & Magnusson, C. (2015). Changes in prevalence of autism spectrum disorders in 2001-2011: Findings from the Stockholm youth cohort. *Journal of Autism and Developmental Disorders*, *45*(6), 1766–1773. https://doi.org/10.1007/s10803-014-2336-y

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. https://doi.org/10.1038/s41380-018-0321-0

Mottron, L. (2021). A radical change in our autism research strategy is needed: back to prototypes. *Autism Research*. https://doi.org/10.1002/aur.2494.

Rødgaard, E. M., Jensen, K., Vergnes, J. N., Soulières, I., & Mottron, L. (2019). Temporal changes in effect sizes of studies comparing individuals with and without autism: A meta-analysis. *JAMA Psychiatry*, *76*(11), 1124–1132. https://doi.org/10.1001/jamapsychiatry.2019.1956

How to cite this article: Mottron, L. (2021). Progress in autism research requires several recognition-definition-investigation cycles. *Autism Research*, 1–5. https://doi.org/10.1002/aur.2524