

Tails or Types? A Critical Systematic Review of Taxometric Studies in Neurodevelopmental Disorders

Enrico Toffalini¹, Margherita Calderan¹, Riccardo Pagan², and Valentina Tobia³

¹Department of General Psychology, University of Padua, Italy

²Department of Social and Developmental Psychology, University of Padua, Italy

³University Vita-Salute San Raffaele, Milan, Italy

Author Note

Enrico Toffalini  <https://orcid.org/0000-0002-1404-5133>

Margherita Calderan  <https://orcid.org/0009-0005-5668-5162>

The review was preregistered at

https://osf.io/65y9g/?view_only=91e97c6b00dc444abb83fefcf67cbac0

https://osf.io/ys5ad/?view_only=0e2812acfabf43fc8af01f883c7331ba Financial support: Project 2022KBW99S “Tails or types? Testing the dimensional hypothesis in neurodevelopmental disorders”, funded by Next Generation EU, Mission 4, Componente 1, CUP C53D23004210006

We are grateful to Athanasia Vasileiadou who first pointed us to some of the reviewed articles.

Correspondence concerning this article should be addressed to Enrico Toffalini, Email: enrico.toffalini@unipd.it

Abstract

...

Keywords: Neurodevelopmental disorders, Taxometric analysis, Autism, ADHD, Learning disorders

Tails or Types? A Critical Systematic Review of Taxometric Studies in Neurodevelopmental Disorders

In recent years, research on neurodevelopmental disorders (NDDs) has undergone widespread reconsiderations. Traditional diagnostic categories defined in the DSM-5 ([American Psychiatric Association, 2013](#)), have been increasingly called into question in favor of a dimensional and transdiagnostic view (e.g., [Astle et al., 2022](#); [Happé & Frith, 2020, 2021](#); [Michelini et al., 2024](#); [Sonuga-Barke, 2020](#)). The DSM-5 itself has incorporated dimensional aspects, for example in defining a broad Autistic Spectrum Disorder (ASD) condition instead of a set of different autism-related diagnoses. Several arguments have contributed to this: the large heterogeneity within diagnostic groups, the poor alignment between data-driven clusters and DSM-based categories, the frequent comorbidity and phenotypic correlations across conditions, and the idea that individuals with NDDs might represent, at least in part, the tails of developmental continua, that are quantitatively rather than qualitatively different from the so-called neurotypical population ([Happé & Frith, 2021](#); [Mammarella et al., 2021](#); [Michelini et al., 2024](#)).

Clarifying whether NDDs truly represent categories, rather than just tails of continuous distributions that span across the population, is important for different reasons. These involve quests about social and personal self-identity by involved individuals, discussions about diagnostic framing, comorbidity, and the choice on interventions in clinical settings, the structuring of classes, schools, and educational methods when dealing with neurodiversity, and applications to research design, planning of studies, and power analysis in research settings ([Caviola et al., 2024](#)). But most importantly, this is relevant for our understanding of the ontological status of reality, which should matter for scientific endeavor.

A dimensional approach fits well with long-standing principles in the study of individual differences in psychology, and has long characterized research on intelligence and personality ([Caviola et al., 2024](#)). It also seems appropriate for NDDs given accumulating evidence of polygenicity and multifactorial etiology ([Astle & Fletcher-Watson, 2020](#); [Caviola et al., 2024](#); [Demontis et al., 2023](#); [Kovas & Plomin, 2006](#); [Pennington et al., 2012](#); [Plomin & Kovas, 2005](#)).

In fact, most behavioral traits are believed to be influenced not by one or a few, but by many small, additive causes ([Chabris et al., 2015](#)). According to a famous principle known in statistics as the central limit theorem ([Fischer, 2011](#)), when many individual causes (even discrete ones, e.g., gene variants) add up they converge to a joint distribution that tends to continuity and normality: this is, in fact, the “secret” of the Gaussian distribution. Nonetheless, a principled categorical view still resists in neurodevelopmental research, for example concerning ASD ([Chown & Leatherland, 2021](#); [Frazier et al., 2023](#); e.g., [Happé & Frith, 2021](#)), subtypes within NDDs (e.g., [Burgess et al., 2018](#)) or the possible emergence of new data-driven taxa ([Astle et al., 2022](#)).

Adjudicating between whether a clinical diagnostic category is characterized by an underlying structure that is categorical or dimensional is precisely the focus of taxometric analysis, a set of quantitative methods that have been developed since the 1990s (e.g., [Meehl, 1995](#)). Since then, taxometric analysis has been widely applied in psychopathology, generally providing evidence that most mental health conditions (e.g., depression, anxiety) are best conceptualized as dimensional rather than categorical ([Haslam et al., 2012](#); [Haslam et al., 2020](#)). In brief, taxometric methods such as MAMBAC, MAXCOV, MAXEIG, and L-Mode, leverage patterns of covariation among observed indicators to test alternative competing latent models that are categorical or dimensional ([Cole, 2004](#); [Ruscio, 2007](#); [Ruscio & Ruscio, 2004](#)). Such methods have become increasingly refined through simulation-based techniques such as the Comparison Curve Fit Index (CCFI), which provide criteria for objective decision-making ([Ruscio et al., 2018](#)).

Despite this, the use of taxometric analysis in neurodevelopmental research has been limited. Most existing studies focus on adult psychopathology, with few exceptions ([Beauchaine, 2003](#); [Haslam et al., 2020](#)). [Happé and Frith \(2021\)](#) and [Chown and Leatherland \(2021\)](#) in their debate on whether ASD is a category or a continuum not even mention the potential of taxometric analysis. In their recent influential article, [Astle et al. \(2022\)](#) briefly presents taxometry but gives it little emphasis, prevailing reviewing studies that employ cluster analysis. However, the limitations of clustering for taxometric purposes have been discussed for decades. Clustering algorithms tend to impose (rather than seek) categorical structure, often fail to recover known taxa

in simulations, and generally provide weak tests of taxonicity ([Beauchaine, 2003](#)). Recent studies simulating realistic psychological data, characterized by a combination of (even modestly) skewed and correlated indicators, further suggest that common clustering algorithms easily detect pseudoclusters while having low power to test the existence of true taxa ([Toffalini et al., 2022, 2024](#)). Indeed, Astle et al. ([2022](#)) interpret results from clustering studies as evidence that data-driven structures fail to align with the expected DSM-5 categories, rather than as proper tests of taxonicity.

Taxometric analysis is not without limitations, however, some of which are shared with cluster analysis. Commonly cited threats to validity of taxometric conclusions include poor indicator validity, nuisance covariance, low base rate, and artificial admixture [...]. Poor indicator validity occurs when taxometric analysis is run on variables that present little mean differences between individuals with and without the target disorder (i.e., putatively, the taxon and the complement). Generally, it is recommended that Cohen's $d \geq 1$ on each indicator. Nuisance covariance means that there is strong residual correlations after accounting for taxa (i.e., both within the taxon and the complement), thus reflecting the use of redundant indicators; residual $r > 0.40$ is often regarded as problematic. Low base rate means that individuals with the taxonic condition are too few in the sample for taxometric analysis to correctly identify them; $n < 30$ is often regarded as problematic. This is obviously a problem when the diagnostic category is rare (e.g., 1-2% of the population). Finally, artificial admixture consists of mixing participants selected or recruited using alternate criteria. This is often (incorrectly) used to compensate for the low base rate, with many individuals with a diagnosis being recruited separately from the rest of the sample to increase the number of observations with the putative taxon. Poor indicator validity and low base rate risk masking detection of taxa, while strong nuisance covariance and artificial admixture risk to detecting pseudotaxa by mimicking categorical structures.

Goals of the Present Review

The main goal of this systematic review is to critically examine the use of taxometric methods in the study of neurodevelopmental disorders (NDDs) within the published literature. We

focused our review on the most commonly studied NDD categories: Autism Spectrum Disorder (ASD), Attention-Deficit/Hyperactivity Disorder (ADHD), and all forms of Specific Learning Disorders (SLDs), along with the closely related condition of Language Disorder/Impairment (the latter was not included in the preregistered protocol and thus constitutes a deviation). We excluded Intellectual Disability, as its etiological basis is often well identified and taxonic by definition in many cases (e.g., Down syndrome), or associated with identifiable medical risk factors. Residual non-syndromic cases, in turn, are thought to reflect the lower tail of general mental ability, and thus as inherently dimensional.

A core aim of the review is to critically assess the methodological quality and credibility of findings reported in published taxometric studies. We evaluated potential sources of bias such as artificial admixture, and whether sufficiently complete methodological information was reported to judge other threats to validity, including low indicator validity, nuisance covariance, and the skewness of indicators. (The latter is relevant when scales assess rare or extreme behaviors, and can create the illusion of categorical structure).

Following this systematic review, we present a Monte Carlo simulation study to empirically demonstrate the extent to which artificial admixture can lead to false detection of categorical latent structures in dimensional data, both when using taxometric analysis and cluster analysis.

Systematic Review of Taxometric Analysis on NDDs

Methods

The systematic review was conducted according to the Preferred Reporting items for Systematic Reviews and Meta-Analysis 2020 guidelines. A review protocol was preregistered on OSF registries at the following link:

https://osf.io/65y9g/?view_only=91e97c6b00dc444abb83fefcf67cbac0 The methods are detailed below.

Sources and Search strategy

A systematic search was conducted for relevant peer-reviewed articles in April 2025 using four electronic databases: PsycINFO, PubMed, Scopus, Web of Science. No time limits were set, although no records were expected from before 1990, as taxometric methods were not yet developed or diffuse then. The following search query, reflecting the methods and the target populations of interest, was performed on titles, abstracts, and keywords:

((“taxometric*” OR “MAMBAC” OR “MAXCOV” OR “MAXEIG” OR “L-Mode”) AND (“autis*” OR “ASD” OR “attention-deficit*” OR “ADHD” OR “learning dis*” OR “reading dis*” OR “dyslexia” OR “math* dis*” OR “math* diff*”))

in databases where the wildcard “*” was not supported (i.e., PubMed), it was replaced with full terms (“disorder”, “disability” instead of “dis*”; “mathematical”, “mathematics”, “math” instead of “math*”).

Language disorder was subsequently added as a deviation from preregistration, replacing the second part of the above search query with (“language imp*” OR “language dis*”). In addition, the full search query was launched on Google Scholar and the first 10 pages were screened; the references of and citations to Haslam et al. (2020) were searched on Google Scholar for additional results; and a deep research via OpenAI’s gpt-4o was launched providing it the preregistration protocol as context. All promising records thus emerging and not previously identified via databases, were listed under the *“Identification of studies via other methods”* section of the PRISMA flowchart.

Eligibility Criteria

Inclusion criteria were: 1) the study applies at least one taxometric procedure (e.g., MAMBAC, MAXCOV, L-Mode, MAXEIG); 2) the target population is individuals with NDDs or related traits; 3) the study draws a conclusion about the latent structure (categorical vs. dimensional) or reports sufficient information for an independent reader to reach a conclusion based on those results (e.g., reports CCFIs).

Exclusion criteria were: 1) the study only reports a theoretical, methodological, or

simulation articles without empirical data; 2) the article does not report any behavioral data (e.g., only neuroimaging or neurophysiological data; this was later reconsidered, but no studies were excluded for this reason in any case); 3) the study only consists of a review and/or meta-analysis; 4) the study employs other latent structure methods only (e.g., LPA, mixture modeling), and not taxometric analysis. In addition, following the preregistration protocol, the search was limited to published, peer-reviewed journal articles, as a primary aim was to provide a critical review of the existing literature. However, one potentially interesting record excluded as non-journal article was separately discussed below in the Results section.

Screening and Coding

Two independent human coders (MC and RP) screened and coded all articles. No AI tools were used in the process, except in occasional cases by the human coders to assist them clarify the content of articles. All discrepancies were resolved through discussion, and later revised by the PI (ET).

Since a single article could report multiple (sub)samples and/or multiple taxometric analyses on the same sample, the full coded dataframe followed a long-form data structure, with some articles repeating over multiple rows. Coded variables were listed in the data dictionary table defined in the preregistration protocol, to which a few variables for notes and coders agreement were added.

Inter-coder agreement was fair to good in all phases. For the title+abstract screening phase, Cohen's $k = 0.79$ (34 out of 38 matching independent decisions); for the full-text eligibility screening, Cohen's $k = 0.52$ (15 out of 19). For final data extraction, 84% (71 out of 85) rows were extracted by both coders, with good agreement on the crucial information: for the "authors' taxonic conclusion" agreement was 94%, for artificial admixture was 95% at the sample/subsample level.

Results

A total of 110 records were identified through database searches. After removing 67 duplicates and excluding 5 records (dissertations or book sections, only one of which appeared to

include primary taxometric analyses, but could not be retrieved), 38 unique articles were screened. Nineteen were excluded at the title/abstract level. Full-text assessment was conducted on the remaining 19 reports, with 7 excluded for not meeting inclusion criteria (4 did not use taxometric analysis; 3 did not target neurodevelopmental disorders). An additional 4 records were identified through citation tracking and online browsing, of which 2 were excluded (1 methodological article; 1 not employing taxometric methods). In total, 14 studies were included in the review. The PRISMA flowchart in Figure 1 summarizes the selection process.

Figure 1 *about here*

An overview summary of the results (simplified to one row per article) is reported in Table 1. A fully detailed reporting of coded data for all taxometric analyses is available in Supplemental online material. The total number of taxometric analyses was 85, spanning from 1 to 22 per article. Targeted disorders were ADHD (6 articles), ASD (6 articles), Dyslexia (1 article), and Specific Language Impairment (SLI; 2 articles), with one article targeting both ADHD and ASD. Most articles (10 out of 14) focused on whether disorders were taxa, while the remaining 4 articles targeted whether subtypes were taxa.

The most used taxometric method was MAMBAC, used in 13 out of 14 articles and in 84 out of 85 total analyses, followed by MAXEIG (9 articles), L-Mode (7 articles), MAXCOV (2 articles), and MAXSLOPE (1 article). In most articles more than one taxometric method was used, with a median of 3 methods per analysis. In 11 out of 14 articles the interpretation was based on a quantitative index (Fit_d or, for more recent articles, CCFI), while in the remaining 3 cases the interpretation was based on visual inspection of taxometric curves. The number of indicators entered in the analyses spanned from 2 to 9, with a median of 3 (excluding one article where the number was unclear). The median sample size across all analyses was 600, or 611.25 after averaging within article.

Taxonic conclusions were mixed. All 5 articles targeting ADHD as a disorder concluded in favor of a dimensional structure, as did both articles targeting SLI. On the contrary, all 4 articles targeting ASD as a disorder leaned towards or clearly supported a categorical (taxonic)

structure. The remaining 4 articles targeted subtypes of disorders, and all leaned towards taxonic conclusions, either strongly or predominantly (Stevens et al., 2018 on ADHD subtypes; Munson et al., 2008 on ASD subtypes; O'Brien et al., 2012 on Dyslexia subtypes), or at least partly (Ingram et al., 2008, that supported ASD subtypes as taxa depending on what sets of indicators were used). In total, 8 out of 14 articles reached taxonic conclusions on at least some of their taxometric analyses (specifically, 28 out of 85 total analyses, or 32.9%, reached a taxonic conclusion), mostly concerning ASD (21 out of 29 analyses, or 72.4%).

Artificial admixture was evident or strongly suspected, for at least part of the analyses, in 8 out of 14 articles (in total, involving 31 out of 85 analyses). Articles where artificial admixture was incurred tended to match those that reached taxonic conclusions, with 6 out of 8 cases. This matching was even higher at the level of individual analysis: taxonic conclusion was reached in 20 out of 31 (64.5%) analyses where admixture was evident or suspected, in contrast to 8 out of 54 (14.8%) analyses where it was probably or certainly excluded.

Remaining methodological issues concerned missing information about descriptive statistics or other indices. In particular, nuisance covariances (i.e., correlations across indicators within diagnostic group or taxa) could be coded for only 2 out of 14 articles. Where reported, it was small to negligible in James et al. (2016), while above 0.40 for many indicators in Frazier et al. (2023); both articles reached a taxonic conclusion. Skewness of distributions of indicators scores was reported by only 5 articles, and where reported it was often close to or beyond ± 1 (at least one indicator in about one third of the analyses); this was not specifically associated to reaching a taxonic conclusion. On the contrary, indicators validity was largely reported, being coded for all but 3 articles, and it was almost always large or very large ($\gg 1$).

Table 1 *about here*

Additional Use of Cluster Analysis

Among the 14 identified articles, 6 also employed methods of cluster analysis (in all but one case using latent class or latent profile analysis). In 5 of the 6 cases, cluster analysis was conducted in parallel with taxometric analysis (Deserno et al., 2023; Frazier et al., 2007; Frazier

et al., 2023; James et al., 2016; Munson et al., 2008), while in one case it was performed as a secondary analysis (Frazier et al., 2010). All 6 targeted ADHD and/or ASD, and in all cases clustering results were interpreted as supporting taxonic conclusions, always converging with those from taxometric analyses, with the exception of Frazier et al. (2007), where clustering results were partially inconclusive yet suggested the presence of subgroups within ADHD, in contrast to the dimensional structure favored by taxometric analysis. In all but 2 studies (Frazier et al., 2007; Munson et al., 2008), we identified at least some risk of artificial admixture.

Simulation of Artificial Admixture

We focus on artificial admixture because it is a very powerful threat to validity that could relatively easily be avoided by design when planning a study, and because as suggested by the result of the systematic review presented below, it is unfortunately frequent. There are different types of admixtures; Ruscio and Ruscio (2004) discusses three cases. First, as mentioned above, the situation where clinical and nonclinical cases are separately recruited and merged. Second, the case of splitting a sample into subsamples (using alternative criteria) and then using those to test different taxometric questions. Third, trimming away observations of the putative complement (i.e., individuals without the disorder) to increase the base rate. All of these procedures artificially impose categorical structures to the data that do not exist in the population. It is worth mentioning that artificial admixture is also a threat for any type of cluster analysis, for the exact same reasons that it is a problem for taxometric analysis.

We simulate a dimensional scenario that makes it visually compelling how artificial admixture of the first type mentioned above creates pseudotaxa. A single normally-distributed latent variable X is generated, that defines the target condition. A series of observed indicators $x_1, x_2, \text{etc.}$ are simulated, incorporating the X latent variable plus some random normally-distributed noise ($R^2 = 0.77$). Individuals above the 99th percentile of X receive a diagnosis via clinical assessment. A sample of $N = 1000$ individuals is collected from the general population. Since the base rate is low (on average, only about $n = 10$ cases are expected to naturally present a diagnosis), researchers select and recruit an additional $n = 1000$ cases with a

clinical diagnosis from the general population, thus creating artificial admixture. All data simulation is available in the online materials at https://osf.io/ys5ad/?view_only=0e2812acfabf43fc8af01f883c7331ba

An instance of the simulation, plotting two observed variables for illustrative purposes, is shown in Figure 2. Color is added for further clarity to distinguish individuals with and without a diagnosis, but two clusters still clearly emerge. Indicator validity is good (about *Cohen's d* = 2), while nuisance covariance is partly problematic: for the clinical cases $r \approx 0$, but for the nonclinical cases $r = 0.61$. It is worth reminding that no taxa exist here, and the appearance of a categorical structure is mainly imposed by admixture, as shown below.

Figure 2 *about here*

A Monte Carlo simulation was then performed, with 1000 iterations. Taxometric analysis was conducted using the “RTaxometric” package of R. As recommended, CCFI was used for inference: CCFI varies between 0 (maximal evidence for dimensional structure) to 1 (maximal evidence for categorical structure), with 0.5 indicating inconclusive results. Performed procedures were, as per default: MAMBAC, MAXEIG, and L-Mode. A third $x3$ observed indicator (similar to those above) was entered to allow conducting all these taxometric procedures. Over 1000 simulated datasets, results always strongly favored a categorical structure: median combined CCFI = 0.88, 95% quantile interval [0.83, 0.92]; median MAMBAC CCFI = 0.90; average MAXEIG CCFI = 0.86; average L-Mode CCFI = 0.88. The smallest combined CCFI over 1000 iterations was 0.80. Therefore, admixture can be an extremely strong driver of false categorical findings, especially when many clinical cases are separately recruited and added to the sample.

To confirm that artificial admixture leads to misleading results also when performing clustering, Gaussian Mixture Models (GMM) was also run testing 1 vs 2 components, using the “mclust” package of R. GMM is a type of finite mixture model, similar to latent profile analysis, but more flexible in that it also models indicator covariances. Over 1000 iterations, results always favored the 2-component relative to the 1-component model, with a median $\Delta BIC = -548.67$ in favor of the 2-component model.

To confirm that the categorical structure was incorrectly favored specifically because of artificial admixture, we run another Monte Carlo simulation like the one above, but without admixture. In this case, a single sample of $N = 2000$ observations was drawn at each iteration. Therefore, nuisance covariance remains strong, but there is no admixture. Median combined CCFI = 0.47, 95% quantile interval [0.40, 0.58]; median MAMBAC CCFI = 0.39; average MAXEIG CCFI = 0.49; average L-Mode CCFI = 0.53. Thus, results remain frequently around inconclusive, probably due to nuisance covariance, but do not risk strongly favoring a categorical structure. GMM also consistently favored the 1-component model (over the 2-component model), with an average $\Delta BIC = -30.36$ in favor of the 1-component model.

Discussion

The present review suggests that taxometric analysis has been underused in neurodevelopmental research, and that many articles using it present relevant methodological limitations or incomplete reporting of crucial information. This is unfortunate in light of the unique relevance of taxometric analysis for the current debate on the categorical vs dimensional nature of NDDs ([Astle et al., 2022](#); [Happé & Frith, 2021](#); [Micheline et al., 2024](#)). The underuse of taxometric analysis in this field was expected: of the 183 articles employing taxometric methods on psychopathology identified by Haslam et al. ([2020](#)) only 15 were placed in the “childhood” category (encompassing all types, even conditions such as sleep problems).

Overall, at the article/sample level there were only 2 major mismatches: Marcus et al. ([2012](#)) reached a dimensional conclusions despite incurring admixture by separately recruiting a large number of clinical cases; and Munson et al. ([2008](#)), who reached a taxonic conclusion about ASD subtypes despite no evidence of admixture (all children had received a diagnosis of ASD based on a seemingly uniform clinical assessment).

In the simulated example, artificial admixture was also accompanied by large nuisance covariance, although it was demonstrated that the latter posed a less severe problem to inference. In fact, the coexistence of low nuisance covariance and good indicator validity is practically impossible under a dimensional scenario. This is because an indicator can largely separate

individuals with vs without a diagnosis if it is strongly related to the latent dimension, but this leads to nuisance covariance across valid indicators. Ultimately, however, such impasse might be interpreted as a direct signal that the true latent structure is dimensional.

While the DSM-5 may appear to promote a categorical model of disorders, it does not explicitly commit to a taxonic view. It is often researchers, not diagnostic manuals, who reify clinical categories as natural kinds, despite their originally pragmatic intent.

The fact that observed scores in psychology often deviate from normality (Micceri, 1989) may reflect psychometric characteristics (e.g., skewed indicators when observing times, accuracies, bounded sum scores) rather than the true shape of the underlying latent traits.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. (5th edition). American Psychiatric Association.
- Astle, D. E., & Fletcher-Watson, S. (2020). Beyond the Core-Deficit Hypothesis in Developmental Disorders. *Current Directions in Psychological Science*, 29(5), 431–437. <https://doi.org/10.1177/0963721420925518>
- Astle, D. E., Holmes, J., Kievit, R., & Gathercole, S. E. (2022). Annual Research Review: The transdiagnostic revolution in neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 63(4), 397–417. <https://doi.org/10.1111/jcpp.13481>
- Beauchaine, T. P. (2003). Taxometrics and developmental psychopathology. *Development and Psychopathology*, 15(3), 501–527. <https://doi.org/10.1017/s0954579403000270>
- Burgess, A. P., Witton, C., Shapiro, L., & Talcott, J. B. (2018). *From Subtypes to Taxons: Identifying Distinctive Profiles of Reading Development in Children* (T. Lachmann & T. Weis, Eds.; pp. 213–233). Springer International Publishing. https://doi.org/10.1007/978-3-319-90805-2_11
- Caviola, S., Greiff, S., & Toffalini, E. (2024). Learning disorders and difficulties: From a categorical to a dimensional perspective. *Learning and Individual Differences*, 113, 102490. <https://doi.org/10.1016/j.lindif.2024.102490>

- Chabris, C. F., Lee, J. J., Cesarini, D., Benjamin, D. J., & Laibson, D. I. (2015). The Fourth Law of Behavior Genetics. *Current Directions in Psychological Science*, 24(4), 304–312.
<https://doi.org/10.1177/0963721415580430>
- Chown, N., & Leatherland, J. (2021). Can a Person be ‘A Bit Autistic’? A Response to Francesca Happé and Uta Frith. *Journal of Autism and Developmental Disorders*, 51(2), 749–751.
<https://doi.org/10.1007/s10803-020-04541-0>
- Cole, D. A. (2004). Taxometrics in psychopathology research: an introduction to some of the procedures and related methodological issues. *Journal of Abnormal Psychology*, 113(1), 3–9.
<https://doi.org/10.1037/0021-843X.113.1.3>
- Demontis, D., Walters, G. B., Athanasiadis, G., Walters, R., Therrien, K., Nielsen, T. T., Farajzadeh, L., Voloudakis, G., Bendl, J., Zeng, B., Zhang, W., Grove, J., Als, T. D., Duan, J., Satterstrom, F. K., Bybjerg-Grauholm, J., Bækved-Hansen, M., Gudmundsson, O. O., Magnusson, S. H., ... Børglum, A. D. (2023). Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nature Genetics*, 55(2), 198–208. <https://doi.org/10.1038/s41588-022-01285-8>
- Deserno, M. K., Bathelt, J., Groenman, A. P., & Geurts, H. M. (2023). Probing the overarching continuum theory: data-driven phenotypic clustering of children with ASD or ADHD. *European Child & Adolescent Psychiatry*, 32(10), 1909–1923.
<https://doi.org/10.1007/s00787-022-01986-9>
- Fischer, H. (2011). *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Springer. <https://doi.org/10.1007/978-0-387-87857-7>
- Frazier, T. W., Chetcuti, L., Al-Shaban, F. A., Haslam, N., Ghazal, I., Klingemier, E. W., Aldosari, M., Whitehouse, A. J. O., Youngstrom, E. A., Hardan, A. Y., & Uljarević, M. (2023). Categorical versus dimensional structure of autism spectrum disorder: A multi-method investigation. *JCPP Advances*, 3(2), e12142. <https://doi.org/10.1002/jcv2.12142>
- Frazier, T. W., Youngstrom, E. A., & Naugle, R. I. (2007). The latent structure of attention-deficit/hyperactivity disorder in a clinic-referred sample. *Neuropsychology*, 21(1),

45–64. <https://doi.org/10.1037/0894-4105.21.1.45>

Frazier, T. W., Youngstrom, E. A., Sinclair, L., Kubu, C. S., Law, P., Rezai, A., Constantino, J. N., & Eng, C. (2010). Autism Spectrum Disorders as a Qualitatively Distinct Category From Typical Behavior in a Large, Clinically Ascertained Sample. *Assessment*, 17(3), 308–320. <https://doi.org/10.1177/1073191109356534>

Happé, F., & Frith, U. (2020). Annual Research Review: Looking back to look forward – changes in the concept of autism and implications for future research. *Journal of Child Psychology and Psychiatry*, 61(3), 218–232. <https://doi.org/10.1111/jcpp.13176>

Happé, F., & Frith, U. (2021). Dimensional or Categorical Approaches to Autism? Both are Needed. A Reply to Nick Chown and Julia Leatherland. *Journal of Autism and Developmental Disorders*, 51(2), 752–753. <https://doi.org/10.1007/s10803-020-04728-5>

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychological Medicine*, 42(5), 903–920. <https://doi.org/10.1017/S0033291711001966>

Haslam, N., McGrath, M. J., Viechtbauer, W., & Kuppens, P. (2020). Dimensions over categories: a meta-analysis of taxometric research. *Psychological Medicine*, 50(9), 1418–1432. <https://doi.org/10.1017/S003329172000183X>

Ingram, D. G., Takahashi, T. N., & Miles, J. H. (2008). Defining autism subgroups: a taxometric solution. *Journal of Autism and Developmental Disorders*, 38(5), 950–960. <https://doi.org/10.1007/s10803-007-0469-y>

James, R. J. E., Dubey, I., Smith, D., Ropar, D., & Tunney, R. J. (2016). The Latent Structure of Autistic Traits: A Taxometric, Latent Class and Latent Profile Analysis of the Adult Autism Spectrum Quotient. *Journal of Autism and Developmental Disorders*, 46(12), 3712–3728. <https://doi.org/10.1007/s10803-016-2897-z>

Kovas, Y., & Plomin, R. (2006). Generalist genes: implications for the cognitive sciences. *Trends in Cognitive Sciences*, 10(5), 198–203. <https://doi.org/10.1016/j.tics.2006.03.001>

Mammarella, I. C., Toffalini, E., Caviola, S., Colling, L., & Szűcs, D. (2021). No evidence for a

- core deficit in developmental dyscalculia or mathematical learning disabilities. *Journal of Child Psychology and Psychiatry*, 62(6), 704–714. <https://doi.org/10.1111/jcpp.13397>
- Marcus, D. K., Norris, A. L., & Coccaro, E. F. (2012). The latent structure of attention deficit/hyperactivity disorder in an adult sample. *Journal of Psychiatric Research*, 46(6), 782–789.
- Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist*, 50(4), 266–275.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105(1), 156–166. <https://doi.org/10.1037/0033-2909.105.1.156>
- Micheline, G., Carlisi, C. O., Eaton, N. R., Elison, J. T., Haltigan, J. D., Kotov, R., Krueger, R. F., Latzman, R. D., Li, J. J., Levin-Aspenson, H. F., Salum, G. A., South, S. C., Stanton, K., Waldman, I. D., & Wilson, S. (2024). Where do neurodevelopmental conditions fit in transdiagnostic psychiatric frameworks? Incorporating a new neurodevelopmental spectrum. *World Psychiatry*, 23(3), 333–357. <https://doi.org/10.1002/wps.21225>
- Munson, J., Dawson, G., Sterling, L., Beauchaine, T., Zhou, A., Koehler, E., Lord, C., Rogers, S., Sigman, M., Estes, A., & Abbott, R. (2008). Evidence for latent classes of IQ in young children with autism spectrum disorder. *American Journal on Mental Retardation*, 113(6), 439–452. <https://doi.org/10.1352/2008.113:439-452>
- O'Brien, B. A., Wolf, M., & Lovett, M. W. (2012). A Taxometric Investigation of Developmental Dyslexia Subtypes. *Dyslexia*, 18(1), 16–39. <https://doi.org/10.1002/dys.1431>
- Pennington, B. F., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., Leopold, D., Samuelsson, S., Byrne, B., Willcutt, E. G., & Olson, R. K. (2012). Individual prediction of dyslexia by single vs. Multiple deficit models. *Journal of Abnormal Psychology*, 121(1), 212–224. <https://doi.org/10.1037/a0025823>
- Plomin, R., & Kovas, Y. (2005). Generalist Genes and Learning Disabilities. *Psychological Bulletin*, 131(4), 592–617. <https://doi.org/10.1037/0033-2909.131.4.592>
- Ruscio, J. (2007). Taxometric Analysis: An Empirically Grounded Approach to Implementing the

Method. *Criminal Justice and Behavior*, 34(12), 1588–1622.

<https://doi.org/10.1177/0093854807307027>

Ruscio, J., Carney, L. M., Dever, L., Pliskin, M., & Wang, S. B. (2018). Using the comparison curve fit index (CCFI) in taxometric analyses: Averaging curves, standard errors, and CCFI profiles. *Psychological Assessment*, 30(6), 744–754. <https://doi.org/10.1037/pas0000522>

Ruscio, J., & Ruscio, A. M. (2004). A conceptual and methodological checklist for conducting a taxometric investigation. *Behavior Therapy*, 35(2), 403–447.

[https://doi.org/10.1016/S0005-7894\(04\)80044-3](https://doi.org/10.1016/S0005-7894(04)80044-3)

Sonuga-Barke, E. J. S. (2020). Editorial: ‘People get ready’: Are mental disorder diagnostics ripe for a Kuhnian revolution? *Journal of Child Psychology and Psychiatry*, 61(1), 1–3.

<https://doi.org/10.1111/jcpp.13181>

Stevens, M. C., Pearlson, G. D., Calhoun, V. D., & Bessette, K. L. (2018). Functional neuroimaging evidence for distinct neurobiological pathways in attention-deficit/hyperactivity disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(8), 675–685.

<https://doi.org/10.1016/j.bpsc.2017.09.005>

Toffalini, E., Gambarota, F., Perugini, A., Girardi, P., Tobia, V., Altoè, G., Giofrè, D., Team, P. C., & Feraco, T. (2024). Clusters that are not there: An R tutorial and a Shiny app to quantify a priori inferential risks when using clustering methods. *International Journal of Psychology*, 59(6), 1183–1198. <https://doi.org/10.1002/ijop.13246>

Toffalini, E., Girardi, P., Giofrè, D., & Altoè, G. (2022). Entia Non Sunt Multiplicanda ... Shall I look for clusters in my cognitive data? *PLOS ONE*, 17(6), e0269584.

<https://doi.org/10.1371/journal.pone.0269584>

Table 1*Overview summary of the reviewed articles.”*

Study	Target disorder	Population analyzed	Age	Taxometric focus	Taxonic conclusion	Artificial admixture	Sample size	No. of taxometric analyses	Taxometric methods	Interpret. based on	No. of indicators	Indicators skewn.	Nuisance covar.	Indicators validity
O'Brien et al. (2012)	Dyslexia	Clinical	children	subtypes	mostly yes	partly yes	517 to 671	6	MAMBAC; MAXEIG; L-Mode	CCFI	3 to 4	unreported	unreported	0.82 to 2.37
Marcus et al. (2012)	ADHD	Mixed clinical and non-clinical	adults	disorder	no	yes	432 to 600	6	MAMBAC; MAXEIG; L-Mode	CCFI	2 to 9	unreported	unreported	2.04 to 3.26
Haslam et al. (2006)	ADHD	General population	children	disorder	no	no	1222 to 1774	2	MAXEIG; MAMBAC	Fid	3	0.45 to 1.49	unreported	2.47 to 3.29
Frazier et al. (2007)	ADHD	Clinical referrals	children to adults	disorder	no	no	191 to 394	22	MAMBAC; MAXEIG; L-Mode	Fid	3 to 9	-1.09 to 0.47	unreported	1.03 to 3.59
Marcus and Barry (2011)	ADHD	General population	children	disorder	no	no	667 to 1078	15	MAMBAC; MAXEIG; L-Mode	CCFI	2 to 9	0.74 to 3.91	unreported	1.96 to 3.26
Ingram et al. (2008)	ASD	Clinical	children	subtypes	partly yes	no	481 to 2254	7	MAMBAC; MAXCOV	CCFI	unclear	unreported	unreported	1.51 to 2.55
Deserno et al. (2023)	ADHD / ASD	Mixed clinical and non-clinical	children	disorder(s)	no for ADHD, mostly yes for ASD	probably yes	203 to 434	4	L-Mode; MAMBAC; MAXEIG	CCFI	3	unreported	unreported	1.86 to 2.74
Frazier et al. (2010)	ASD	Clinical and unaffected siblings	children	disorder	yes	partly yes	1825 to 11472	6	MAMBAC; MAXEIG	CCFI	3 to 4	unreported	unreported	unreported
James et al. (2016)	ASD	General population	adults	disorder	yes	yes	1139	1	MAMBAC; MAXCOV; MAXEIG; L-Mode	CCFI	8	-0.48 to 0.73	-0.09 to 0.28	0.15 to 2.21
Munson et al. (2008)	ASD	Clinical	children	subtypes	yes	no	456	1	MAXCOV	visual	4	unreported	unreported	unreported
Stevens et al. (2018)	ADHD	Mixed clinical and non-clinical	adolescents	subtypes	yes	partly yes	71 to 251	2	MAMBAC; MAXS-LOPE	CCFI	3	unreported	unreported	unreported
Frazier et al. (2023)	ASD	Mixed clinical and non-clinical	children to adults	disorder	yes	yes	512 to 16755	10	MAMBAC; MAXEIG; L-Mode	CCFI	3 to 8	0.19 to 0.99	0.14 to 0.73	0.57 to 1.84
Dollaghan (2004)	SLI	General population	children	disorder	no	no	620 to 623	2	MAMBAC	visual	2	unreported	unreported	1.93 to 2.81
Dollaghan (2011)	SLI	General population	children	disorder	no	no	601	1	MAMBAC	visual	2	unreported	unreported	3.47 to 3.52

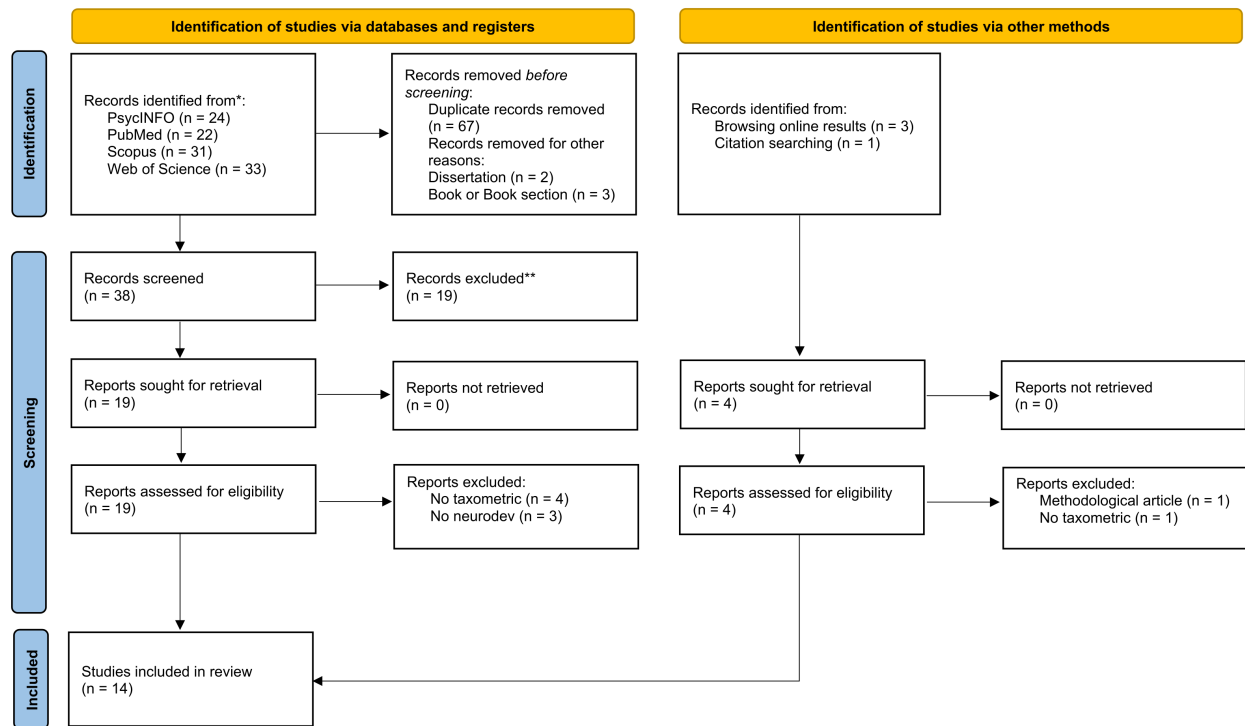
Figure 1*PRISMA 2020 flow diagram illustrating the study screening and selection process*

Figure 2

A simulated example of artificial admixture, where a large number of individuals with a diagnosis are separately recruited and added to the sample (total $N = 2000$). The true underlying structure is strictly dimensional.

