ARTICLE IN PRESS

Child Abuse & Neglect xxx (xxxx) xxx

Contents lists available at ScienceDirect

Child Abuse & Neglect

journal homepage: www.elsevier.com/locate/chiabuneg

Leveraging multivariate approaches to advance the science of early-life adversity

Alexis Brieant [a,b,*], Lucinda M. Sisk [b], Taylor J. Keding [b], Emily M. Cohodes [b], Dylan G. Gee [b]

[a] University of Vermont, Department of Psychological Science, 2 Colchester Avenue, Burlington, VT 05402, USA
[b] Yale University, Department of Psychology, 100 College Street, New Haven, CT 06510, USA

ARTICLE INFO

Keywords:
Adversity
Development
Multivariate
Methodology

ABSTRACT

Since the landmark Adverse Childhood Experiences (ACEs) study, adversity research has expanded to more precisely account for the multifaceted nature of adverse experiences. The complex data structures and interrelated nature of adversity data require robust multivariate statistical methods, and recent methodological and statistical innovations have facilitated advancements in research on childhood adversity. Here, we provide an overview of a subset of multivariate methods that we believe hold particular promise for advancing the field's understanding of early-life adversity, and discuss how these approaches can be practically applied to explore different research questions. This review covers data-driven or unsupervised approaches (including dimensionality reduction and person-centered clustering/subtype identification) as well as supervised/prediction-based approaches (including linear and tree-based models and neural networks). For each, we highlight studies that have effectively applied the method to provide novel insight into early-life adversity. Taken together, we hope this review serves as a resource to adversity researchers looking to expand upon the cumulative approach described in the original ACEs study, thereby advancing the field's understanding of the complexity of adversity and related developmental consequences.

1. Introduction

The landmark Adverse Childhood Experiences (ACEs) study demonstrated a clear, graded relation between exposure to adversity during childhood and physical and mental health outcomes in adulthood (Felitti et al., 1998). These findings spurred a widespread public health response, as well as a body of scientific research that sought to better understand the consequences of adversity for human development. The original study categorized ACEs as abuse (psychological, physical, and sexual) and household dysfunction (substance use, mental illness, domestic violence, and incarceration), operationalized as a summed score of the total number of exposures. In the 25 years following this study, adversity research has expanded to more precisely account for the multifaceted and interrelated nature of adverse experiences, with increased emphasis on the importance of factors such as developmental timing, chronicity, and severity, as well as consideration of forms of adversity that were not captured in the original ACEs measure (e.g., perceptions of identity-based discrimination and racial trauma; Bernard et al., 2020; Cohodes et al., 2021; Manly et al., 2001).

Researchers have increasingly accounted for these factors in data acquisition and analysis, yielding novel insights into how, why,

* Corresponding author at: 338 John Dewey Hall, Department of Psychological Science, University of Vermont, Burlington, VT 05402, USA.
E-mail address: alexis.brieant@uvm.edu (A. Brieant).

https://doi.org/10.1016/j.chiabu.2024.106754
Received 25 October 2023; Received in revised form 12 February 2024; Accepted 14 March 2024

and when adversity impacts child development (Gee, 2021). Furthermore, the increasing availability of large, publicly accessible datasets and archives (e.g., National Data Archive on Child Abuse and Neglect [NDACAN]; Adolescent Brain Cognitive Development [ABCD] Study) provide rich data characterizing a wide range of child experiences, which can facilitate these lines of inquiry. However, these data also introduce challenges in how to best operationalize and organize adversity data. Adversity data are often heterogeneous, non-normally distributed, have complex data structures (e.g., combinations of continuous and categorical data), and are interrelated in nature (e.g., high rates of co-occurrence across domains of adversity). These features necessitate robust *multivariate*[1] statistical methods that can account for this complexity and more accurately represent youths' lived experiences.

Methodological and statistical innovations have facilitated advancements in child adversity research broadly (see the 2019 special issue in *Child Abuse and Neglect* on methods and measurement; Gabrielli & Jackson, 2019). Here, we describe a subset of multivariate methods that we believe hold particular promise for advancing the field's understanding of early-life adversity, given the often complex and varied nature of adversity exposure data. Such data have become more common in recent years—the operationalization of adversity exposure, and *dimensions* of adversity along which effects are hypothesized to cleave, has been the subject of numerous theoretical models (Callaghan & Tottenham, 2016; Cohodes et al., 2021; Ellis et al., 2022; McEwen, 1998; McLaughlin et al., 2021; Smith & Pollak, 2020). Given this diversity of approaches in adversity measurement and operationalization, in the present paper we aim not to suggest a particular theoretical approach to studying adversity, but rather to recognize that burgeoning theoretical complexity necessitates a more comprehensive suite of analysis tools, many of which are capable of handling non-Gaussian (i.e., non-normal) data. Thus, the goal of the present paper is to 1) provide an overview of prevailing multivariate methods that may be especially beneficial for advancing adversity research, and 2) highlight examples of studies that have effectively applied these methods to provide novel insight into early-life adversity. We do not intend to provide an in-depth tutorial on applications of each methodological approach (though we do point to methodological and statistical papers that provide greater depth for readers interested in applying these approaches); rather, we seek to provide adversity researchers with an overview of some of the quantitative tools available to them and how these tools can be practically applied to explore research questions of interest. We have organized these methods into several key domains: data-driven or *unsupervised* approaches (including *dimensionality reduction* and person-centered *clustering*/subtype identification) and *supervised*/prediction-based approaches (including linear and tree-based models and neural networks). These methods are summarized visually in Fig. 1, and definitions for these approaches and related technical terms (italicized in text) are summarized in a glossary in the supplemental material.

2. Unsupervised/data-driven approaches

In contrast to theory-driven approaches that test specific hypotheses, unsupervised, or data-driven, methods are approaches that are often exploratory in nature and that generate insights based on naturally-occurring patterns in a dataset (Maass et al., 2018). Data-driven approaches aid in parsing complexity raised by co-occurrence and, further, relax constraints that may be imposed by existing theory, researcher bias, or preconceptions (Jack et al., 2018), facilitating novel theory-building to advance understanding of the complexity of adversity and related developmental consequences. In the following sections, we provide an overview of several unsupervised approaches with notable applications to adversity research.

2.1. Dimensionality reduction

Dimensionality reduction methods reduce the number of *features* in a dataset, often aiming to reduce measurement noise, account for feature dependency, and/or limit *overfitting* during model building. In theory, these reduced dimensions are thought to represent the latent structure of variables, and serve as a way of reducing *high dimensional* datasets into a subset of variables that account for the majority of variance. This is especially beneficial when datasets have a large number of correlated variables, as is often the case with adversity data. The following methods facilitate identification of patterns and can clarify the relations among different variables by grouping them in statistically meaningful ways.

2.1.1. Exploratory factor analysis

Exploratory factor analysis (*EFA*) is one approach to dimensionality reduction that serves to identify a set of latent (unobserved) factors thought to be the underlying cause of multiple manifest (observed) variables based on their interrelationships (Krishnakumar & Nagar, 2008; Matsunaga, 2010). EFA is, by definition, exploratory in nature and is not, in and of itself, suited for hypothesis testing. However, it can be helpful in characterizing the data's structure, especially with multiple indicators of the same underlying constructs. This can help inform novel theoretical conceptualizations that can then be formally tested in subsequent work.

As one example, Brieant et al. (2023) tested an EFA based on 60 early-life adversity variables in the ABCD Study in order to examine covariation among a wide range of adversities. Estimating the EFA in a structural equation modeling framework (Asparouhov & Muthén, 2009; Marsh et al., 2014) helped to accommodate the complex data structure (i.e., combination of dichotomous, polytomous, and continuous data) across multiple measurement methods, compare model fit across different factor solutions, and adjust for skewed variable distributions—all factors that are especially beneficial when working with multidimensional adversity data. Results indicated 10 robust dimensions of early-life adversity co-occurrence, corresponding to conceptual domains such as caregiver substance use and

[1] italicized terms are defined in the Glossary in Appendix A

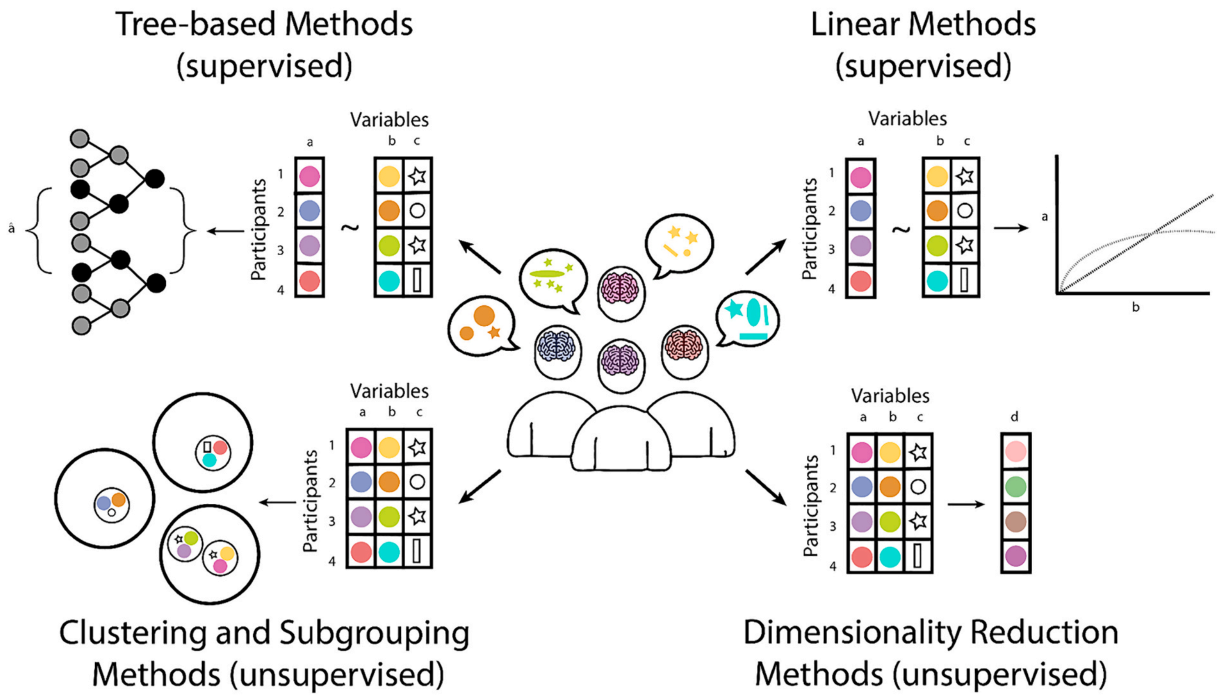

Fig. 1. High-level graphical depiction of four major methodological categories described in the present paper. We distinguish between linear regression and tree-based methods, which characterize associations *between* measurements, clustering and subgrouping methods, which characterize shared patterns *across* measurements, and dimensionality reduction methods, which describe latent structure *underlying* measurements. Importantly, not all methods fall neatly into these groups, and the statistical assumptions and designed purpose of each method vary.

biological caregiver separation, caregiver psychopathology, caregiver lack of support, and socioeconomic disadvantage/neighborhood lack of safety.[2] Importantly, these factors demonstrated predictive validity in that they were associated with distinct behavioral correlates, including internalizing and externalizing symptoms and cognitive functioning (Brieant et al., 2023). Though studies like this illustrate the utility of applying an EFA approach to clarify interrelations among types of adversity (see Table 1 for additional examples), it has been suggested that the causal structure of *reflective models* may not be well-suited for adversity research (e.g., Bethell et al., 2017; McLaughlin et al., 2023). That is, it may not always be the case that there is a "common cause" for the shared variance between different adversity exposures (see Plamondon et al., 2022 for further discussion). Depending on the research question and nature of the data, *formative models* (such as principal component analysis) may be more appropriate, as described below.

2.1.2. Principal component analysis

Similar to EFA, *principal component analysis* (*PCA*) is a dimensionality reduction approach that summarizes and reduces the overall number of variables in a dataset; however, unlike EFA (in which latent factors are modeled to be the cause of the measured indicators), in PCA the underlying dimensions (or "components") are modeled to be the effect of the measured indicators. This results in uncorrelated (or orthogonal) components that maximize explained variance in the original variables (Bryant & Yarnold, 1995; Krishnakumar & Nagar, 2008). Importantly, in contrast to EFA which attempts to differentiate error and signal covariances, PCA does not distinguish between variance due to measurement error versus true signal when estimating components. As such, PCA may not be an appropriate approach with datasets suspected of containing unreliable or high-noise measurements.

As one example, Nikolaidis et al. (2022) applied PCA to a set of caregiving-related early adversity variables, including physical abuse, parental mental illness, and emotional neglect. Fourteen total items were distilled into two significant components which were interpreted as 1) additive exposure (which captured the majority of adversity items) and 2) emotional maltreatment without domestic violence. This approach facilitates a more parsimonious organization of data by reducing dimensionality while simultaneously minimizing loss of information from the original variables. Additional adversity-related PCA examples that further illustrate these benefits are listed in Table 1.

[2] Code for the models from Brieant et al. (2023) is publicly available via the Open Science Foundation: https://osf.io/28cb7/?view_only=b7789e2eb92d40d290358a5ad623ac65. Factor scores for the 10 adversity factors are available for use by researchers with an active data use agreement for the ABCD Study.

Table 1
Summary of outlined multivariate methods.

Method	Goal	Important considerations	Key methodological citations	Examples of application in adversity research	Suggested software/packages
Exploratory Factor Analysis	**Dimensionality reduction** - characterize dimensional structure of variables, identify latent factors	Exploratory, not suited for hypothesis testing; latent factors thought to be the underlying cause of multiple manifest variables based on their interrelationships	Krishnakumar & Nagar, 2008; Matsunaga, 2010	Brieant et al., 2023; Ford et al., 2014; Nikolaidis et al., 2022; Scott et al., 2013	*psych* package in R (Revelle, 2023); Mplus, EFA and ESEM options (Asparouhov & Muthén, 2009)
Principal Components Analysis	**Dimensionality reduction** - reduce the number of variables	Underlying dimensions (or “components”) are modeled to be the *effect* of the measured indicators	Bryant & Yarnold, 1995; Krishnakumar & Nagar, 2008	Nikolaidis et al., 2022; Nweze et al., 2023	*Stats* package in R (R Core Team, 2012); *FactorMineR* in R (Lê et al., 2008); *Scikit-learn* in python (Abraham et al., 2014)
Canonical Correlation Analysis	**Dimensionality reduction** - identify latent “variates” that represent combinations of variables	Enables the identification of latent links between two high dimensional datasets; has linear and non-linear model types	Hotelling, 1936; Kettenring, 1971	Alnæs et al., 2020; Modabbernia et al., 2021; Sisk et al., 2023	*PMD* in R (González et al., 2008; Witten et al., 2009); *cca-zoo* in python (Chapman and Wang, 2021)
Data Fusion	**Dimensionality reduction** - identify shared variance across two or more datasets or data types	Describes various different methods, such as Joint and Individual Variation Explained (JIVE) and Similarity Network Fusion (SNF)	Wang et al., 2014; Yu et al., 2017	Hong & Sisk et al., 2021	*R.jive* in R (O’Connell and Lock, 2016); *SNFtool* in R (Wang et al., 2014)
K-Means	**Person or Variable Clustering** - create subgroups based on similarity between cases using a distance measure	Minimizes the sum of variances within each subgroup	Hartigan & Wong, 1979	Basu et al., 2023; Ehrlich et al., 2016	*Stats* package in R (R Core Team, 2012); *Scikit-learn* in python (Abraham et al., 2014)
Spectral Clustering	**Person or Variable Clustering** - create subgroups based on similarity between cases using a distance measure	Uses graph theory to identify similarities between cases	Newman, 2006; Yu & Shi, 2003	Hong & Sisk et al., 2021; Sheridan et al., 2020	*Spectrum* in R (John et al., 2020); *Scikit-learn* in python (Abraham et al., 2014)
Latent Class Analysis and Latent Profile Analysis	**Person Clustering** - identify latent subgroups of individuals based on response patterns on a set of items	Model-based, can evaluate fit to the data; can account for classification uncertainty. Can also distinguish residual variance/noise from true signal.	Bauer & Shanahan, 2007; Lanza & Cooper, 2016; Nylund-Gibson & Choi, 2018; Sinha et al., 2021	Brieant et al., 2022; Brown et al., 2019; Conley et al., 2022; King et al., 2019; Lian et al., 2022; Pears et al., 2008	Mplus, LPA and LCA options (Muthen & Muthen, 2000); *tidyLPA* in R (Rosenberg et al., 2019)
Regularized Linear Models	**Supervised Prediction** - learn a set of linear parameters with a loss function penalty	Fewest implementation barriers of discussed models. May achieve feature selection (L_1), parameter grouping (L_2) or both (elastic net). Less likely to overfit than non-regularized.	Arnold et al., 2020; Melkumova & Shatskikh, 2017; Schölkopf, 2000; Ying, 2019; Zou & Hastie, 2005	Herd et al., 2022; Keding et al., 2021	*LASSO*, *Ridge*, and *ElasticNet* from Scikit-Learn in Python (Pedregosa et al., 2011); *glmnet* in R (Friedman et al., 2010)
Support Vector Machines	**Supervised Prediction** - learn a decision hyperplane based on maximizing or minimizing its margin from support vectors	Few implementation barriers. Robust to small sample sizes and low computational complexity. Sensitive to outliers and overfitting with high dimensional data.	Cortes & Vapnik, 1995; Shawe-Taylor & Sun, 2011	Carmel & Widom, 2020; Keding et al., 2021	*SVC* and *SVR* from Scikit-Learn in Python (Pedregosa et al., 2011); Interface *libsvm* and package *e1071* in R (Chang & Lin, 2011)
Random Forest	**Supervised Prediction** - learn a set of independently-fit decision trees from bootstrapped subsets of data and aggregate their predictions	Highly-transparent decisions. Makes few/no assumptions about data and parameter structure. Reduced likelihood of overfitting and moderately robust to small samples.	Breiman, 2001; Bühlmann, 2012; Kern et al., 2019	Keding et al., 2021; Teicher et al., 2018	*RandomForestClassifier* and *RandomForestRegressor* from Scikit-Learn in Python (Pedregosa et al., 2011)
Gradient Boosting Trees	**Supervised Prediction** - learn a set of sequentially-fit decision trees, each learning from the prediction errors of previous (weaker) models	Highly-transparent decisions. Makes few/no assumptions about data and parameter structure. Reduced likelihood of overfitting and moderately robust to small samples.	Bühlmann, 2012; Friedman, 2001; Natekin & Knoll, 2013	Khan et al., 2015; Silveira et al., 2020	*GradientBoostingClassifier* and *GradientBoostingRegressor* from Scikit-Learn in Python (Pedregosa et al., 2011); *XGBoost* (Chen & Guestrin, 2016)
Multilayer Perceptron (Neural Networks)	**Supervised Prediction** - learn a set of weights and biases belonging to neurons that interact with each other within interconnected layers	Best for problems where variable interactions are important. Least transparent model discussed. Highest implementation barriers. Requires larger sample sizes and most likely to overfit without regularization, cross validation, etc.	Bebis & Georgiopoulos, 1994; Hanson & Burr, 1990; Liu et al., 2017; Smys et al., 2020	Keding & Herringa, 2022; Keding et al., 2021	*MLPClassifier* and *MLPRegressor* from SciKit-Learn (Pedregosa et al., 2011); *Tensorflow* (Abadi et al., 2016); *PyTorch* (Paszke et al., 2019); all in Python

2.1.3. Canonical correlation analysis

Canonical correlation analysis (*CCA*) comprises a suite of unsupervised, multivariate statistical models that enable the simultaneous identification of latent links between two high dimensional datasets. That is, CCA seeks to identify latent "variates" that represent optimal linear combinations of variables from two independent datasets. CCA is somewhat analogous to PCA, in that variables "load" onto a CCA variate in a similar manner to PCA component. However, unlike PCA, in CCA pairs of variates are identified that maximally correlate with each other, thereby identifying latent links across datasets. Further differences exist between CCA and PCA in the mathematical computations underpinning latent variable identification, and indeed PCA can be used as a dimensionality reduction step prior to CCA (e.g. Alnæs et al., 2020; Smith et al., 2015). Pairs of correlated CCA variates (each variate representing variable loadings from its respective dataset) comprise a "mode". CCA models can fit multiple modes per model, and these modes are predominantly orthogonal to each other. In practice, these properties make CCA particularly well-suited for exploratory research questions aiming to characterize meaningful patterns of shared variance between high dimensional datasets. After identifying modes whose variates are correlated at levels above chance, researchers may then examine the variable loadings on these variates to identify which variables comprise them. Several different types of CCA models have been developed over the years to suit different questions and data structures. For instance, penalized or sparse CCA models (e.g., Witten et al., 2009) have the added benefit of handling *multicollinear* data more efficiently and parsimoniously (Wang et al., 2020; Witten et al., 2009). Additional types of CCA models include multiset CCA (Nielsen, 2002), which can handle more than two high dimensional datasets; non-linear or kernel CCA, which can identify complex nonlinear associations; and probabilistic (or Bayesian) CCA. However, it is important to note that CCA is best suited for moderately-sized to large datasets, and requires careful consideration when computing statistical significance and controlling for nuisance covariates (Winkler et al., 2020).

Although CCA was first introduced in the 1930s (Hotelling, 1936), its high computational demand precluded it from being widely used until recently. With the advent of modern computing systems, CCA has begun gaining popularity in neuroscience and psychology research alongside the development of software packages that make running complex models more feasible. Several papers highlight the promise of CCA for neuroimaging research in particular (Wang et al., 2020; Zhuang et al., 2020), given the high dimensionality and covariance common in neuroimaging datasets. Applications in the context of adversity have been limited to date; however, sparse CCA (sCCA; Witten et al., 2009) was recently applied to identify how adversity exposures at different ages across childhood and adolescence were linked with patterns of white matter tract integrity across three measures (generalized fractional anisotropy, radial diffusivity, and quantitative anisotropy; Sisk et al., 2023). Three sCCA models were fitted—one for each of the three white matter tract integrity measures. The variates comprising the first mode for each of the three models were examined, revealing that across all models exposure to adversity in middle childhood (particularly during ages 5–6 and 8–9) was linked with alterations in white matter tract integrity at levels significantly above chance. In particular, white matter tracts demonstrated a divergent pattern of loadings in relation to middle childhood adversity exposure, such that white matter tracts broadly supporting sensorimotor information (such as the corticospinal tract) loaded opposite white matter tracts broadly supporting higher-order cognitive and emotion processing (such as the superior longitudinal fasciculus). This pattern suggests that adversity during middle childhood may be related to alterations in white matter integrity across the whole brain, and that the directionality of links between adversity and tract-level integrity may differ as a function of the neural circuits the tract supports. Further, the mean loadings for white matter tract integrity in the quantitative anisotropy model were associated with trauma-related symptoms, suggesting that latent whole-brain patterns of tract integrity may relate to mental health symptoms above and beyond measures from a single tract. Given the flexible, data-driven nature of CCA, it is well-suited to research questions that seek to apply a data-driven approach to identify multidimensional links between adversity and neural, behavioral, or clinical measures.

2.1.4. Data fusion

Data fusion approaches represent another promising method for researchers aiming to apply data-driven approaches across multiple types of data, particularly when datasets include variables of different types (e.g., numeric, categorical) and on different scales. Across scientific domains, many different approaches have been developed in service of identifying shared signal, minimizing noise, and more comprehensively characterizing a participant sample across two or more data types (Castanedo, 2013). For *multimodal* research studies, these methods present a promising alternative to common practices such as conducting separate analyses within modalities, or simply concatenating datasets (Wang et al., 2014). Notably, methods previously described in the current paper such as principal components analysis (PCA) and canonical correlation analysis (CCA) have also been described as data fusion (Calhoun & Sui, 2016). While the characterization of and mathematical computations underpinning data fusion methods vary, here we will highlight Similarity Network Fusion (SNF; Wang et al., 2014), which leverages an iterative, nonlinear fusion approach to characterize shared patterns of variation across datasets. Specifically, SNF uses a two-step network approach: first, sample-similarity networks are constructed for each type of data, and second, these networks are fused into a single similarity network using an iterative nonlinear process (Wang et al., 2014). The resulting similarity network has been shown to be scalable, robust to outliers, and able to identify meaningful information even in relatively small sample sizes (Wang et al., 2014). Following data fusion, applying a clustering model to the resulting similarity network may facilitate the identification of subtypes of individuals. One recent study leveraging SNF examined shared variance between brain and environmental data in the ABCD Study, and how these patterns differed across participants (Hong & Sisk et al., 2021). SNF was applied to measures of cortical thickness, myelination, and environmental factors (specifically trauma exposure, neighborhood safety, school environment, and family environment). Subsequently, *spectral clustering* was applied to identify subtypes characterized by shared variation in the similarity network comprising both neural and environmental data. Results demonstrated that two- and five-subtype solutions fit the data best, and demonstrated that measures of key environmental risk and protective factors were differentially linked with patterns of cortical thickness and myelination in youth. Further, incorporating subtyping information into a predictive model facilitated clinical symptom prediction above and beyond imaging or environmental features alone. These results suggest that data fusion may aid in clarifying heterogeneity across individuals and data types, thereby parsimoniously representing individual- and subgroup-level differences in brain, behavior, and mental health.

2.2. Person clustering and subgroup identification

The methods described thus far are often considered "variable-centered" (or "feature-centered"), as they attempt to organize and describe relations between variables. In contrast, clustering methods that are "person-centered" (or "*instance*-centered") offer opportunities to group individuals based on shared attributes (Bauer & Shanahan, 2007). In the context of adversity research, this can be especially helpful in parsing heterogeneity in adverse experiences by identifying subgroups of individuals who share similar types of adversity exposure and who are meaningfully distinct from other subgroups of individuals. In this way, these approaches can characterize specific patterns of adversity exposure and help inform more individualized intervention efforts that are tailored to youths' unique and varied experiences. This can be accomplished through a variety of methods, including K-means and spectral clustering methods, latent class analysis, and latent profile analysis.

2.2.1. K-means and spectral clustering

Certain clustering methods use a distance measure to create subgroups based on similarity between cases. Optimally, cases that are most similar belong to the same subgroup, while cases that are least similar belong to different subgroups. Identified subgroups are often highly dependent on the choice of distance measure (e.g., Euclidean distance, Manhattan distance, cosine similarity). Additionally, different clustering algorithms such as *K-Means* (Hartigan & Wong, 1979) and *spectral clustering* (Newman, 2006; Yu & Shi, 2003) use different logic to identify subgroups. Thus, the choice of which algorithm, as well as which distance measure, is most suitable is dependent on the specific research question and type of data available (see Shirkhorshidi et al., 2015; Xu & Tian, 2015). For instance, K-Means clustering minimizes the sum of variances within each subgroup, iteratively computing this *metric* and reassigning subgroups until minimum values for each subgroup are obtained (Morissette & Chartier, 2013). In contrast, spectral clustering uses graph theory to identify similarities between cases, indexed as low *edge* weights (i.e., relatively weaker connections) between cases in different subgroups, which are iteratively computed and subgroups reassigned until minimum values for each subgroup are obtained (von Luxburg, 2007).

As one example, a recent study examined whether a clustering approach would yield subtypes of individuals differentially characterized by dimensions of adversity exposure (i.e., threat and deprivation) and cognitive and emotional abilities (Sheridan et al., 2020). Association networks were constructed, representing correlation strengths between each possible pairing of variables while controlling for every other variable. Next, four clustering methods were applied and consensus clustering (Lancichinetti & Fortunato, 2012) was used to determine the most representative groupings. Across two separate datasets, the resulting clusters suggested that dimensions of adversity may be differentially linked with cognitive and emotional abilities in youth.

2.2.2. Latent class and latent profile analysis

Latent class analysis (*LCA*) is a model-based clustering approach that can be used to identify unobserved subgroups of individuals in a structural equation modeling (SEM) framework (Nylund-Gibson & Choi, 2018). Generally speaking, LCA is distinct from other clustering methods such as K-Means in that it derives clusters using a probabilistic model, rather than a distance measure. Thus, relative goodness-of-fit can be evaluated for the models and guide the determination of the most appropriate number of clusters. Posterior probabilities are also generated, which allows assessment of the certainty of class membership. For these reasons, LCA is often considered a more robust clustering method, although it is more computationally demanding (Sinha et al., 2021). The LCA approach assumes the presence of multiple mutually-exclusive "classes" within a population, characterized by different patterns of responses on a set of categorical variables (Lanza & Cooper, 2016) and has increasingly been adopted in adversity research. As one example, Brown et al. (2019) applied LCA to identify patterns of maltreatment among a large sample of children who had contact with child welfare services. Based on dichotomous reports of ACEs exposure, three distinct clusters of exposure emerged (with some variation depending on developmental stage): 1) physical neglect, emotional abuse, and domestic violence; 2) physical neglect/household dysfunction; and 3) emotional abuse. This study nicely illustrates how LCA can offer insight into complex, co-occurring experiences within individuals; however, the reliance of LCA on categorical data may omit important information about the degree or severity of adversity exposure. Latent profile analysis (LPA) is computationally similar to LCA but uses continuous, rather than categorical, indicators to characterize subgroups. For example, Brieant et al. (2022) estimated latent profiles in a community sample of adolescents based on eight different continuous variables, including factors such as exposure to abuse, neglect, household chaos, and socioeconomic disadvantage. Results indicated four different adversity profiles: low exposure, neglect, household instability, and poly-adversity. Importantly, individuals with these profiles significantly differed on mental health outcomes four years later; the poly-adversity group reported significantly higher levels of both internalizing and externalizing symptoms relative to the low exposure group, and the household instability group showed elevated risk for externalizing symptoms. Linking subgroups with outcomes of interest in this way can provide a form of validation that the subgroups are substantively meaningful, while also providing important insight into the varied developmental consequences of adversity. Additional LCA and LPA examples are listed in Table 1.

3. Supervised/prediction-based approaches

The methods discussed thus far have been predominantly unsupervised methods, wherein the model does not have access to "ground truth" expectations. In contrast, supervised methods incorporate an external source of knowledge into model fitting, using the difference between model predictions and *supervisory signals* to train the model. Prediction-based approaches typically use supervised learning to create a model that maps the relationship between a set of input variables (referred to as features) and the supervisory signal and shows low *generalization error* on unobserved data. Broadly, prediction-based methods have three components that must be decided by the researcher a priori: the model *algorithm*, *loss function*, and the *optimizer*. Here, the algorithm is the class of model being implemented (e.g., linear, tree-based), the loss function refers to the equation used to calculate the model's prediction errors, or deviations from the supervisory signal (e.g., mean squared error, cross entropy), and the optimizer is the method by which model parameters are updated (after exposure to data) with the aim of minimizing or maximizing the loss function (e.g., maximum likelihood estimation, stochastic gradient descent). In the following sections, we provide an overview of common algorithms, important considerations for their appropriate use, and their application to early-life adversity research.

3.1. Linear models

Linear[3] models are one of the most common types of prediction-based approaches, primarily due to their ease of implementation, low *computational complexity*, and interpretability (Arnold et al., 2020). Linear models assume that the supervisory signal can be predicted by a weighted sum of the input features; importantly, these features can take nonlinear forms, including interactions (multiplication of two or more features) and other transformation functions (e.g., log transform). Additionally, linear models are extremely adaptable to a variety of data types, ranging from nominal- to ratio-level measurements, for both the supervisory signal (through *link functions*; e.g., logistic for binary outcomes, Poisson for count outcomes) and features. Modern computers can optimize linear models quickly relative to other models, allowing them to scale well in situations with large samples and a large number of features. At the same time, because the weighted-sum form of linear models is less complex than other models, they are also less prone to overfitting and can show adequate fit with smaller sample sizes. For these reasons, linear models represent a good starting point, or "default choice", of a model algorithm, often performing well in a wide variety of situations.

Although linear models can scale well to many features, there are cases where the number of features overwhelms the available sample size ($p > n$) or where many features are correlated with each other, causing issues related to multicollinearity. In these cases, *regularized* linear models can help overcome such challenges during model fitting. Regularization refers to a *penalty* added to the model's loss function, increasing the model's loss as the number of parameters and their magnitudes increase. These penalties have the effect of decreasing model complexity and reducing overfitting (Ying, 2019). For example, the *least absolute shrinkage and selection operator* (LASSO; L_1 penalty) iteratively updates low-magnitude parameters toward 0; this produces a model that implements *feature selection*. Another example is *ridge* (L_2 penalty) regularization, which reduces overfitting through a "grouping" effect on related parameters; here, the parameters of correlated variables will tend to increase and decrease together, preventing arbitrary feature dropout observed with LASSO and limiting multicollinearity issues. Additionally, ridge regularization penalizes extreme feature values more harshly, making the model more robust to outliers (see Melkumova & Shatskikh, 2017 for more information).

Elastic net regularization (Zou & Hastie, 2005) utilizes both the L_1 and L_2 penalties in the same loss function. This combination allows the model to implement feature selection like L_1 while being less sensitive to multicollinearity and more robust to outliers like L_2. For example, Herd et al. (2022) used a combination of subgroup identification and prediction-based approaches to study how early-life adversity was related to different trajectories of post-traumatic stress symptoms (PTSS) in female adolescents. First, they identified three distinct trajectories of PTSS: (1) recovery; (2) moderate, chronic; and (3) high, chronic; second, an elastic net model tested how adversity could predict trajectory membership. Their results demonstrated that the recovery trajectory was characterized by the absence of sexual abuse, physical abuse, and other traumas, less affective dysregulation, having fewer high risk-taking friends, lower levels of parent depression, and being of racial/ethnic minority status (Herd et al., 2022). Feature selection and decreased sensitivity to multicollinearity were particularly important here, allowing Herd et al. to overcome the challenge of including many, highly-related adversity features in a single model and discovering environmental markers of recovery from PTSS.

Finally, *support vector machines* (SVMs; see Shawe-Taylor & Sun, 2011) are traditionally considered linear models, although nonlinear implementations do exist. Broadly, SVMs learn parameters that define a *hyperplane* that maximally separates data points belonging to different categories given some *margin* (i.e., the "thickness" of the hyperplane). Here, the boundary of the margin is defined by specific data points known as *support vectors*. Like other linear models, various implementations of SVMs support different types of measurement for inputs and supervisory signals, and can be used for both classification (nominal-level) and regression (ratio-level) analyses. SVMs also perform well in datasets with smaller sample sizes; however, hyperplanes can be sensitive to outliers, especially in cases of high dimensionality where the likelihood that a support vector contains an outlier increases. Carmel and Widom (2020) used SVMs to develop an adult-retrospective measure of severe early-life neglect to overcome analytic challenges due to missing exposure reports (especially early in development) and degraded memory about such exposures over time. SVMs were used to predict the optimal exposure items to include in an overall propensity score for experiencing early-life neglect. The final measure consisted of 10 items, which successfully predicted adults who suffered severe neglect and may not have been identified or treated at the time of exposure (Carmel & Widom, 2020). Propensity scores demonstrated strong predictive, construct, and discriminant validity, highlighting strengths of using prediction-based approaches.

[3] "Linear" does not limit these algorithms to only detecting linear relationships, but instead refers to linearity of the parameters, where each must take the form $\beta X + C$ (but variable X can be transformed with any nonlinear function).

3.2. Tree-based models

Another category of prediction-based models, tree-based algorithms, are *nonparametric* (in contrast to *parametric* models), indicating that the structure of the model is learned as part of the model fitting process itself. Here, a sequence of decision points, based on thresholded feature values, is learned from the data; these decision points, when diagrammed, take the form of a directed tree, where the first *node* (the root) is theoretically the "best" feature to estimate groups or values and the remaining nodes predict with greater and greater specificity (see Kern et al., 2019). Because features are "split" at specific values at each decision point of the tree, these models are a useful choice when interactions between features are theoretically important or predictive of the supervisory signal. Additionally, these models are known to perform exceptionally well with tabular data as input, where the structural organization of the features is unrelated to the values those features take (i.e., where there is no information contained in the spatial/temporal relationship between features, as is the case with image or time-series data). Finally, such models are robust against datasets with small sample sizes and can be sensitive to high dimensionality.

Random forest models (Breiman, 2001)—which utilize *decision tree ensembles,* as well as bootstrap aggregation (*bagging*; Bühlmann, 2012)—are a powerful extension of decision trees. To implement random forest models, a bootstrapped sample of the original data is generated and a random subsample of candidate features is selected; this sample is subsequently used to fit a single decision tree. This is repeated *n* times (the number of trees in the forest, specified a priori), and predictions from the trees are aggregated to form a final prediction. The combination of ensemble and bagging approaches make random forests robust to overfitting and outliers, but when taken together, protect against poorly-performing submodels. Teicher et al. (2018) used random forests combined with feature importance analysis to assess the most important predictors of adult hippocampal volume among multiplicity, severity, duration, and/or timing of maltreatment exposures at different periods of development. They discovered that male hippocampal volume was predicted by neglect, but not abuse, experiences up through 7 years of age. Female hippocampal volume, on the other hand, was predicted by abuse, but not neglect, experiences at 10, 11, 15 and 16 years. Interestingly, exposure at peak age had greater predictive importance than multiplicity, severity, or duration (Teicher et al., 2018).

Another tree-based algorithm is *gradient-boosting trees* (or gradient-boosting machines [GBMs]; Friedman, 2001). GBMs are similar to random forest models, utilizing ensemble and bagging approaches to combine multiple decision tree models into a larger model. However, unlike random forests, whose underlying decision trees are fit independently, GBMs attempt to predict the errors of previous trees in the ensemble (see Natekin & Knoll, 2013 for more information). This type of ensemble fitting is known as *boosting*. Silveira et al. compared different tree-based prediction methods to investigate the links between childhood trauma, executive dysfunction, and brain functional connectivity in predicting high-risk drinking in adolescents; they showed childhood trauma severity was significantly related to executive dysfunction across all developmental milestones. At baseline, functional connectivity patterns in regions like the dorsal anterior cingulate cortex, anterior insula, intraparietal sulcus, and postcentral gyri mediated the relationship between childhood trauma and executive dysfunction (Silveira et al., 2020).

3.3. Neural networks

Artificial neural networks (ANNs) are a connectionist (Hanson & Burr, 1990) subclass of *machine learning*, often referred to as *deep learning*. ANN models consist of nodes/*neurons*, representing the features in a dataset, which are organized into layers and are interconnected across layers through edges. Like tree-based algorithms, ANNs are ideal for situations where feature interactions are crucial to predicting the supervisory signal; however, unlike tree-based models, ANNs are uniquely suited to analyzing data where the structural organization of the features is related to the values those features take (i.e., where important information is contained in the spatial/temporal relationship between features), as is the case with images (e.g., computer vision), time series (e.g. audio processing, natural language processing), or a combination of the two (e.g., video processing, neuroimaging analysis). Indeed, research on ANNs has produced many unique network architectures (see Liu et al., 2017; Smys et al., 2020), each designed to handle specific needs of various data structures. The backbone of these architectures is the *multilayer perceptron* (MLP; Bebis & Georgiopoulos, 1994). An MLP's network structure represents a latent, interactive feature map (a collection of nodes/neurons), making MLPs ideal for modeling systems with complex interactions (see the Universal Approximation Theorem; Cybenko, 1989; Pinkus, 1999). However, it is because of this property that MLPs are extremely sensitive to overfitting, especially with small samples, and thus *hyperparameter tuning*, regularization, and cross validation are almost always necessary.

ANNs and their derivative architectures have unique challenges that limit their use, most of which involve their complexity and *black box* nature. These typically include being more difficult to implement, debug, tune, and train, requiring extensive computational resources as models get larger. Additionally, because of the complex structure of ANNs, models typically have many (often by orders of magnitude) more parameters to be estimated relative to other models, necessitating large sample sizes to ensure generalizability; however, with an appropriately-chosen architecture and large sample, ANNs tend to out-perform most other model algorithms in their predictive performance. As such, their use is only beginning to emerge in adversity research. For example, Keding et al. used an ensemble learning approach to combine predictions from many algorithms discussed here, including ridge, SVM, random forest, GBM, and MLP regression models; here, they aimed to predict chronological age from gray matter brain features in female children and adolescents with a history of physical and/or sexual abuse. Results indicated that female participants exposed to abuse showed delayed maturation of emotion-related neural circuits relative to those not exposed to abuse. This delay was observed regardless of subsequent expression of internalizing psychopathology and was associated with altered structural maturation of key subcortical brain structures involved in fear learning, including the hippocampus (Keding et al., 2021). As large, population-representative datasets become commonplace in adversity research, neural network approaches are likely to become more common, as their ability to capture complex, interactive patterns supersedes other approaches.

4. Discussion

Each of the described methods can be leveraged to advance the field's understanding of adversity in different ways. We have focused on a subset of multivariate, quantitative approaches that hold particular promise in the context of adversity research given their ability to handle common challenges in adversity data, such as multicollinearity, heterogeneity, skewed distributions, and complex data structures. While the present paper is by no means exhaustive, and there are numerous other methods with applications in adversity research that are not covered here due to space constraints (e.g., graph theory, network analysis), we hope that this review serves as a starting point for adversity researchers interested in expanding upon the methodology of the original ACEs study. Leveraging such powerful modeling approaches holds great promise for advancing the field's understanding of the complex, interrelated, and multidimensional nature of adversity and its developmental consequences.

4.1. Interplay between unsupervised/data-driven methods and theory

Theory-driven hypothesis testing has remained the prevailing approach in adversity research. Here, we have reviewed the ways in which a selection of unsupervised/data-driven approaches can be implemented to advance this area of research in complementary ways; indeed, these perspectives are synergistic and are both essential to the scientific process. Translating causal, mechanism-driven theories of adversity into precise, testable, and falsifiable hypotheses has proven challenging (due in large part to the complexity and heterogeneity of adversity) and the literature to date lacks consensus (Smith & Pollak, 2020). Thus, identifying data-driven patterns of co-occurrence or covariance that are clinically relevant (i.e., that map onto mental health-related variables) represents an important, complementary branch of research that can parallel theory-driven explorations and further facilitate hypothesis generation and refinement of existing theory. Importantly, exploratory approaches are not entirely atheoretical, as variable selection is often guided by existing theory, and they also allow for inclusion of novel factors that may not be represented in existing theory (Van Lissa, 2022). For example, factors such as the predictability and controllability of experiences have historically not been explicitly accounted for in theoretical models of adversity; however, these factors are important elements of youths' experiences (Baram et al., 2012; Cohodes et al., 2021; Ellis et al., 2022; Smith & Pollak, 2020). Multivariate methods can facilitate understanding of how these features of adversity exposure (among others) interrelate, co-occur, and differentially predict youth outcomes.

4.2. Optimizing multivariate approaches

In order to maximize the benefits of these types of multivariate approaches, there are key considerations that researchers can account for during study conceptualization and data acquisition. The original ACEs questionnaire (Felitti et al., 1998) included a small number of items related to abuse and household dysfunction, and participants reported on whether or not they had experienced those particular adversities as children. Notably, there is no clear rationale delineated in the original ACEs study for the selection of these particular adversities (Lacey & Minnis, 2019). Following the ACEs study, more comprehensive and nuanced measures of adversity have been developed, and refined measurement of adverse experiences is critical to further advancing this area of research. For example, the Maltreatment and Abuse Chronology of Exposure (MACE; Teicher & Parigger, 2015) assesses a wide range of features including specific types of abuse and neglect, timing of exposure, and severity. The Dimensional Inventory of Stress and Trauma across the Lifespan (DISTAL; Cohodes et al., 2023) provides further expansion by querying features such as controllability, predictability, discrimination, and physical proximity to a stressor. The multidimensional, heterogeneous data derived from these measures are well-suited for advanced multivariate methods that can accommodate the large number of interrelated variables and clarify their relative importance for youth outcomes.

4.3. Conclusion

The original ACEs study by Felitti and colleagues was crucial in establishing the robust relations between adversity and long-term health and well-being. Twenty-five years later, significant advances in quantitative methods have ushered forth a new wave of adversity research that has the potential to fundamentally shift the field's conceptualization of adversity. Advanced multivariate statistical approaches can help address some of the key challenges in adversity research and facilitate improvements in intervention and prevention efforts that mitigate the consequences of adversity for children and adolescents.

Funding

A.B. is supported by a National Institutes of Health (NIH) Loan Repayment Program award (L40MH131146); L.M.S. by a National Science Foundation Graduate Research Fellowship Program award (NSF DGE-1752134); E.M.C. by a National Science Foundation Graduate Research Fellowship Program award (NSF DGE1752134), The Society for Clinical Child and Adolescent Psychology (Division 53 of the American Psychological Association) Donald Routh Dissertation Grant, the American Psychological Foundation Elizabeth Munsterberg Koppitz Child Psychology Graduate Fellowship, a Dissertation Funding Award from the Society for Research in Child Development, a Dissertation Research Award from the American Psychological Association; and D.G.G. by a National Science Foundation (NSF) CAREER award (BCS-2145372).

CRediT authorship contribution statement

Alexis Brieant: Writing – original draft, Writing – review & editing, Conceptualization. **Lucinda M. Sisk:** Conceptualization, Visualization, Writing – original draft, Writing – review & editing. **Taylor J. Keding:** Conceptualization, Writing – original draft, Writing – review & editing. **Emily M. Cohodes:** Conceptualization, Writing – review & editing. **Dylan G. Gee:** Conceptualization, Writing – review & editing.

Declaration of competing interest

None.

Data availability

No data was used for the research described in the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chiabu.2024.106754.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., … Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. In *Proceedings of the 12th USENIX conference on operating systems design and implementation* (pp. 265–283).

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., … Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics, 8*. https://doi.org/10.3389/fninf.2014.00014

Alnæs, D., Kaufmann, T., Marquand, A. F., Smith, S. M., & Westlye, L. T. (2020). Patterns of sociocognitive stratification and perinatal risk in the child brain. *Proceedings of the National Academy of Sciences, 117*, 12419–12427. https://doi.org/10.1073/pnas.2001517117

Arnold, K. F., Davies, V., de Kamps, M., Tennant, P. W. G., Mbotwa, J., & Gilthorpe, M. S. (2020). Reflection on modern methods: Generalized linear models for prognosis and intervention—Theory, practice and implications for machine learning. *International Journal of Epidemiology, 49*(6), 2074–2082. https://doi.org/10.1093/ije/dyaa049

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Baram, T. Z., Davis, E. P., Obenaus, A., Sandman, C. A., Small, S. L., Solodkin, A., & Stern, H. (2012). Fragmentation and unpredictability of early-life experience in mental disorders. *American Journal of Psychiatry, 169*(9), 907–915. https://doi.org/10.1176/appi.ajp.2012.11091347

Basu, D., Ghosh, A., Naskar, C., Balachander, S., Fernandes, G., Vaidya, N., … Benegal, V. (2023). Risk clustering and psychopathology from a multi-center cohort of Indian children, adolescents, and young adults. *Development and Psychopathology, 35*(2), 800–808. https://doi.org/10.1017/S0954579422000050

Bauer, D. J., & Shanahan, M. J. (2007). Modeling complex interactions: Person-centered and variable-centered approaches. In *Modeling contextual effects in longitudinal studies* (pp. 255–283). Lawrence Erlbaum Associates Publishers.

Bebis, G., & Georgiopoulos, M. (1994). Feed-forward neural networks. *IEEE Potentials, 13*(4), 27–31. https://doi.org/10.1109/45.329294

Bernard, D. L., Calhoun, C. D., Banks, D. E., Halliday, C. A., Hughes-Halbert, C., & Danielson, C. K. (2020). Making the "C-ACE" for a culturally-informed adverse childhood experiences framework to understand the pervasive mental health impact of racism on black youth. *Journal of Child & Adolescent Trauma*. https://doi.org/10.1007/s40653-020-00319-9

Bethell, C. D., Carle, A., Hudziak, J., Gombojav, N., Powers, K., Wade, R., & Braveman, P. (2017). Methods to assess adverse childhood experiences of children and families: Toward approaches to promote child well-being in policy and practice. *Academic Pediatrics, 17*(7 Suppl), S51–S69. https://doi.org/10.1016/j.acap.2017.04.161

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brieant, A., Clinchard, C., Deater-Deckard, K., Lee, J., King-Casas, B., & Kim-Spoon, J. (2022). Differential associations of adversity profiles with adolescent cognitive control and psychopathology. *Research on Child and Adolescent Psychopathology*. https://doi.org/10.1007/s10802-022-00972-8

Brieant, A., Vannucci, A., Nakua, H., Harris, J., Lovell, J., Brundavanam, D., … Gee, D. G. (2023). Characterizing the dimensional structure of early-life adversity in the Adolescent Brain Cognitive Development (ABCD) study. *Developmental Cognitive Neuroscience, 61*, Article 101256. https://doi.org/10.1016/j.dcn.2023.101256

Brown, S. M., Rienks, S., McCrae, J. S., & Watamura, S. E. (2019). The co-occurrence of adverse childhood experiences among children investigated for child maltreatment: A latent class analysis. *Child Abuse & Neglect, 87*, 18–27. https://doi.org/10.1016/j.chiabu.2017.11.010

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In *Reading and understanding multivariate statistics* (pp. 99–136). American Psychological Association.

Bühlmann, P. (2012). Bagging, boosting and ensemble methods. In J. E. Gentle, W. K. Härdle, & Y. Mori (Eds.), *Handbook of computational statistics: Concepts and methods* (pp. 985–1022). Springer. https://doi.org/10.1007/978-3-642-21551-3_33.

Calhoun, V. D., & Sui, J. (2016). Multimodal fusion of brain imaging data: A key to finding the missing link(s) in complex mental illness. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*(3), 230–244. https://doi.org/10.1016/j.bpsc.2015.12.005

Callaghan, B. L., & Tottenham, N. (2016). The stress acceleration hypothesis: Effects of early-life adversity on emotion circuits and behavior. *Current Opinion in Behavioral Sciences, 7*, 76–81. https://doi.org/10.1016/j.cobeha.2015.11.018

Carmel, T., & Widom, C. S. (2020). Development and validation of a retrospective self-report measure of childhood neglect. *Child Abuse & Neglect, 106*, Article 104555. https://doi.org/10.1016/j.chiabu.2020.104555

Castanedo, F. (2013). A review of data fusion techniques. *TheScientificWorldJournal, 2013*, Article 704504. https://doi.org/10.1155/2013/704504

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, *2*(3), 27:1–27:27. doi: https://doi.org/10.1145/1961189.1961199.

Chapman, J., & Wang, H.-T. (2021). CCA-Zoo: A collection of Regularized, Deep Learning based, Kernel, and Probabilistic CCA methods in a scikit-learn style framework. *Journal of Open Source Software, 6*(68), 3823. https://doi.org/10.21105/joss.03823

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Cohodes, E. M., Kitt, E. R., Baskin-Sommers, A., & Gee, D. G. (2021). Influences of early-life stress on frontolimbic circuitry: Harnessing a dimensional approach to elucidate the effects of heterogeneity in stress exposure. *Developmental Psychobiology*. https://doi.org/10.1002/dev.21969

Cohodes, E. M., McCauley, S., Pierre, J. C., Hodges, H. R., Haberman, J. T., Santiuste, I., … Gee, D. G. (2023). Development and validation of the Dimensional Inventory of Stress and Trauma Across the Lifespan (DISTAL): A novel assessment tool to facilitate the dimensional study of psychobiological sequelae of exposure to adversity. *Developmental Psychobiology, 65*(4), Article e22372. https://doi.org/10.1002/dev.22372

Conley, M. I., Hernandez, J., Salvati, J. M., Gee, D. G., & Baskin-Sommers, A. (2022). The role of perceived threats on mental health, social, and neurocognitive youth outcomes: A multicontextual, person-centered approach. *Development and Psychopathology, 1–22*. https://doi.org/10.1017/S095457942100184X

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297. https://doi.org/10.1007/BF00994018

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems, 2*(4), 303–314. https://doi.org/10.1007/BF02551274

Ehrlich, K. B., Ross, K. M., Chen, E., & Miller, G. E. (2016). Testing the biological embedding hypothesis: Is early life adversity associated with a later proinflammatory phenotype? *Development and Psychopathology, 28*(4pt2), 1273–1283. https://doi.org/10.1017/S0954579416000845

Ellis, B. J., Sheridan, M. A., Belsky, J., & McLaughlin, K. A. (2022). Why and how does early adversity influence development? Toward an integrated model of dimensions of environmental experience. *Development and Psychopathology, 1–25*. https://doi.org/10.1017/S0954579421001838

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., … Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) study. *American Journal of Preventive Medicine, 14*(4), 245–258. https://doi.org/10.1016/S0749-3797(98)00017-8

Ford, D. C., Merrick, M. T., Parks, S. E., Breiding, M. J., Gilbert, L. K., Edwards, V. J., … Thompson, W. W. (2014). Examination of the factorial structure of adverse childhood experiences and recommendations for three subscale scores. *Psychology of Violence, 4*(4), 432–444. https://doi.org/10.1037/a0037723

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*(1), 1–22.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gabrielli, J., & Jackson, Y. (2019). Innovative methodological and statistical approaches to the study of child maltreatment: Introduction. *Child Abuse & Neglect, 87*, 1–4. https://doi.org/10.1016/j.chiabu.2018.12.001

Gee, D. G. (2021). Early adversity and development: Parsing heterogeneity and identifying pathways of risk and resilience. *American Journal of Psychiatry, 178*(11), 998–1013. https://doi.org/10.1176/appi.ajp.2021.21090944

González, I., Déjean, S., Martin, P. G. P., & Baccini, A. (2008). CCA: An R package to extend canonical correlation analysis. *Journal of Statistical Software, 23*, 1–14. https://doi.org/10.18637/jss.v023.i12

Hanson, S. J., & Burr, D. J. (1990). What connectionist models learn: Learning and representation in connectionist networks. *Behavioral and Brain Sciences, 13*(3), 471–489. https://doi.org/10.1017/S0140525X00079760

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 28*(1), 100–108. https://doi.org/10.2307/2346830

Herd, T., Haag, A.-C., Selin, C., Palmer, L., S., S., Strong-Jones, S., … Noll, J. G. (2022). Individual and social risk and protective factors as predictors of trajectories of post-traumatic stress symptoms in adolescents. *Research on Child and Adolescent Psychopathology.*. https://doi.org/10.1007/s10802-022-00960-y

Hong, S.-J., Sisk, L. M., Caballero, C., Mekhanik, A., Roy, A. K., Milham, M. P., & Gee, D. G. (2021). Decomposing complex links between the childhood environment and brain structure in school-aged youth. *Developmental Cognitive Neuroscience, 48*, Article 100919. https://doi.org/10.1016/j.dcn.2021.100919

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika, 28*(3/4), 321–377. https://doi.org/10.2307/2333955

Jack, R. E., Crivelli, C., & Wheatley, T. (2018). Data-driven methods to diversify knowledge of human psychology. *Trends in Cognitive Sciences, 22*(1), 1–5. https://doi.org/10.1016/j.tics.2017.10.002

John, C. R., Watson, D., Barnes, M. R., Pitzalis, C., & Lewis, M. J. (2020). Spectrum: Fast density-aware spectral clustering for single and multi-omic data. *Bioinformatics, 36*(4), 1159–1166. https://doi.org/10.1093/bioinformatics/btz704

Keding, T., & Herringa, R. (2022). Violence exposure and amygdala-prefrontal circuit maturation: Developmental markers of psychiatric risk in youth. *Biological Psychiatry, 91*(9), S47–S48. https://doi.org/10.1016/j.biopsych.2022.02.140

Keding, T. J., Heyn, S. A., Russell, J. D., Zhu, X., Cisler, J., McLaughlin, K. A., & Herringa, R. J. (2021). Differential patterns of delayed emotion circuit maturation in abused girls with and without internalizing psychopathology. *The American Journal of Psychiatry, 178*(11), 1026–1036. https://doi.org/10.1176/appi.ajp.2021.20081192

Kern, C., Klausch, T., & Kreuter, F. (2019). Tree-based machine learning methods for survey research. *Survey Research Methods, 13*(1), 73–93.

Kettenring, J. R. (1971). Canonical analysis of several sets of variables. *Biometrika, 58*(3), 433–451. https://doi.org/10.1093/biomet/58.3.433

Khan, A., McCormack, H. C., Bolger, E. A., McGreenery, C. E., Vitaliano, G., Polcari, A., & Teicher, M. H. (2015). Childhood maltreatment, depression, and suicidal ideation: Critical importance of parental and peer emotional abuse during developmental sensitive periods in males and females. *Frontiers in Psychiatry, 6*. https://www.frontiersin.org/articles/10.3389/fpsyt.2015.00042.

King, L. S., Humphreys, K. L., Camacho, M. C., & Gotlib, I. H. (2019). A person-centered approach to the assessment of early life stress: Associations with the volume of stress-sensitive brain regions in early adolescence. *Development and Psychopathology, 31*(2), 643–655. https://doi.org/10.1017/S0954579418000184

Krishnakumar, J., & Nagar, A. L. (2008). On exact statistical properties of multidimensional indices based on principal components, factor analysis, MIMIC and structural equation models. *Social Indicators Research, 86*(3), 481–496. https://doi.org/10.1007/s11205-007-9181-8

Lacey, R. E., & Minnis, H. (2019). Practitioner review: Twenty years of research with adverse childhood experience scores – Advantages, disadvantages and applications to practice. *Journal of Child Psychology and Psychiatry.*. https://doi.org/10.1111/jcpp.13135

Lancichinetti, A., & Fortunato, S. (2012). Consensus clustering in complex networks. *Scientific Reports, 2*, 336. https://doi.org/10.1038/srep00336

Lanza, S. T., & Cooper, B. R. (2016). Latent class analysis for developmental research. *Child Development Perspectives, 10*(1), 59–64. https://doi.org/10.1111/cdep.12163

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software, 25*, 1–18. https://doi.org/10.18637/jss.v025.i01

Lian, J., Kiely, K. M., & Anstey, K. J. (2022). Cumulative risk, factor analysis, and latent class analysis of childhood adversity data in a nationally representative sample. *Child Abuse & Neglect, 125*, Article 105486. https://doi.org/10.1016/j.chiabu.2022.105486

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing, 234*, 11–26. https://doi.org/10.1016/j.neucom.2016.12.038

Maass, W., Parsons, J., Purao, S., Storey, V. C., & Woo, C. (2018). Data-driven meets theory-driven research in the era of big data: Opportunities and challenges for information systems research. *Journal of the Association for Information Systems, 1253–1273*. https://doi.org/10.17705/1jais.00526

Manly, J. T., Kim, J. E., Rogosch, F. A., & Cicchetti, D. (2001). *Dimensions of child maltreatment and children's adjustment: Contributions ofdevelopmental timing and subtype* (p. 24).

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*(1), 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research, 3*(1), 97–110. https://doi.org/10.21500/20112084.854

McEwen, B. S. (1998). Stress, adaptation, and disease: Allostasis and allostatic load. *Annals of the New York Academy of Sciences, 840*(1), 33–44. https://doi.org/10.1111/j.1749-6632.1998.tb09546.x

McLaughlin, K. A., Sheridan, M. A., Humphreys, K. L., Belsky, J., & Ellis, B. J. (2021). *The value of dimensional models of early experience: Thinking clearly about concepts and categories. 10*.

McLaughlin, K. A., Weissman, D. G., & Flournoy, J. (2023). Challenges with latent variable approaches to operationalizing dimensions of childhood adversity – A commentary on Sisitsky et al. (2023). *Research on Child and Adolescent Psychopathology*. https://doi.org/10.1007/s10802-023-01114-4

Melkumova, L. E., & Shatskikh, S. Y. (2017). Comparing ridge and LASSO estimators for data analysis. *Procedia Engineering, 201*, 746–755. https://doi.org/10.1016/j.proeng.2017.09.615

Modabbernia, A., Janiri, D., Doucet, G. E., Reichenberg, A., & Frangou, S. (2021). Multivariate patterns of brain-behavior-environment associations in the adolescent brain and cognitive development study. *Biological Psychiatry, 89*(5), 510–520. https://doi.org/10.1016/j.biopsych.2020.08.014

Morissette, L., & Chartier, S. (2013). The k-means clustering technique: General considerations and implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology, 9*(1), 15–24. https://doi.org/10.20982/tqmp.09.1.p015

Muthen, B., & Muthen, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24*(6), 882–891. https://doi.org/10.1111/j.1530-0277.2000.tb02070.x

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics, 7*. https://www.frontiersin.org/articles/10.3389/fnbot.2013.00021.

Newman, M. E. J. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E, 74*(3), Article 036104. https://doi.org/10.1103/PhysRevE.74.036104

Nielsen, A. A. (2002). Multiset canonical correlations analysis and multispectral, truly multitemporal remote sensing data. *IEEE Transactions on Image Processing, 11*(3), 293–305. https://doi.org/10.1109/83.988962

Nikolaidis, A., Heleniak, C., Fields, A., Bloom, P. A., VanTieghem, M., Vannucci, A., … Tottenham, N. (2022). Heterogeneity in caregiving-related early adversity: Creating stable dimensions and subtypes. *Development and Psychopathology, 1–14*. https://doi.org/10.1017/S0954579421001668

Nweze, T., Ezenwa, M., Ajaelu, C., & Okoye, C. (2023). Childhood mental health difficulties mediate the long-term association between early-life adversity at age 3 and poorer cognitive functioning at ages 11 and 14. *Journal of Child Psychology and Psychiatry, 64*(6), 952–965. https://doi.org/10.1111/jcpp.13757

Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science, 4*(4), 440–461. https://doi.org/10.1037/tps0000176

O'Connell, M. J., & Lock, E. F. (2016). R.JIVE for exploration of multi-source molecular data. *Bioinformatics, 32*(18), 2877–2879. https://doi.org/10.1093/bioinformatics/btw324

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., … Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems, 32*. https://proceedings.neurips.cc/paper_files/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html.

Pears, K. C., Kim, H. K., & Fisher, P. A. (2008). Psychosocial and cognitive functioning of children with specific profiles of maltreatment. *Child Abuse & Neglect, 32*(10), 958–971. https://doi.org/10.1016/j.chiabu.2007.12.009

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research, 12*(null), 2825–2830.

Pinkus, A. (1999). Approximation theory of the MLP model in neural networks. *Acta Numerica, 8*, 143–195. https://doi.org/10.1017/S0962492900002919

Plamondon, A., Racine, N., McDonald, S., Tough, S., & Madigan, S. (2022). Disentangling adversity timing and type: Contrasting theories in the context of maternal prenatal physical and mental health using latent formative models. *Development and Psychopathology, 34*(5), 1961–1973. https://doi.org/10.1017/S0954579421000353

R Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Revelle, W. (2023). *psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois: Northwestern University. R package version 2.4.1 https://CRAN.R-project.org/package=psych.

Rosenberg, J. M., Beymer, P. N., Anderson, D. J., Lissa, C.j.v., & Schmidt, J. A. (2019). tidyLPA: An R package to easily carry out latent profile analysis (LPA) using open-source or commercial software. *Journal of Open Source Software, 3*(30), 978. https://doi.org/10.21105/joss.00978

Schölkopf, B. (2000). The kernel trick for distances. *Advances in Neural Information Processing Systems, 13*. https://papers.nips.cc/paper_files/paper/2000/hash/4e87337f366f72daa424dae11df0538c-Abstract.html.

Scott, B. G., Burke, N. J., Weems, C. F., Hellman, J. L., & Carrión, V. G. (2013). The interrelation of adverse childhood experiences within an at-risk pediatric sample. *Journal of Child & Adolescent Trauma, 6*(3), 217–229. https://doi.org/10.1080/19361521.2013.811459

Shawe-Taylor, J., & Sun, S. (2011). A review of optimization methodologies in support vector machines. *Neurocomputing, 74*(17), 3609–3618. https://doi.org/10.1016/j.neucom.2011.06.026

Sheridan, M. A., Shi, F., Miller, A. B., Salhi, C., & McLaughlin, K. A. (2020). Network structure reveals clusters of associations between childhood adversities and development outcomes. *Developmental Science, 23*(5), Article e12934. https://doi.org/10.1111/desc.12934

Shirkhorshidi, A. S., Aghabozorgi, S., & Wah, T. Y. (2015). A comparison study on similarity and dissimilarity measures in clustering continuous data. *PLoS One, 10*(12), Article e0144059. https://doi.org/10.1371/journal.pone.0144059

Silveira, S., Shah, R., Nooner, K. B., Nagel, B. J., Tapert, S. F., de Bellis, M. D., & Mishra, J. (2020). Impact of childhood trauma on executive function in adolescence—Mediating functional brain networks and prediction of high-risk drinking. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 5*(5), 499–509. https://doi.org/10.1016/j.bpsc.2020.01.011

Sinha, P., Calfee, C. S., & Delucchi, K. L. (2021). Practitioner's guide to latent class analysis: Methodological considerations and common pitfalls. *Critical Care Medicine, 49*(1), e63–e79. https://doi.org/10.1097/CCM.0000000000004710

Sisk, L. M., Keding, T. J., Cohodes, E. M., McCauley, S., Pierre, J. C., Odriozola, P., … Gee, D. G. (2023). Multivariate links between the developmental timing of adversity exposure and white matter tract integrity in adulthood [preprint]. *Neuroscience*. https://doi.org/10.1101/2023.11.12.566271

Smith, K. E., & Pollak, S. D. (2020). Rethinking concepts and categories for understanding the neurodevelopmental effects of childhood adversity. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691620920725, 174569162092072.

Smith, S., Nichols, T., Vidaurre, D., Winkler, A., Behrens, T., Glasser, M., Ugurbil, K., Barch, D., Van Essen, D., & Miller, K. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature Neuroscience, 18*(11), 1565–1567. https://doi.org/10.1038/nn.4125

Smys, D. S., Chen, D. J. I. Z., & Shakya, D. S. (2020). Survey on neural network architectures with deep learning. *Journal of Soft Computing Paradigm, 2*(3), 186–194.

Teicher, M. H., Anderson, C. M., Ohashi, K., Khan, A., McGreenery, C. E., Bolger, E. A., … Vitaliano, G. D. (2018). Differential effects of childhood neglect and abuse during sensitive exposure periods on male and female hippocampus. *NeuroImage, 169*, 443–452. https://doi.org/10.1016/j.neuroimage.2017.12.055

Teicher, M. H., & Parigger, A. (2015). The 'Maltreatment and Abuse Chronology of Exposure' (MACE) scale for the retrospective assessment of abuse and neglect during development. *PLoS One, 10*(2), Article e0117423. https://doi.org/10.1371/journal.pone.0117423

Van Lissa, C. J. (2022). Developmental data science: How machine learning can advance theory formation in Developmental Psychology. *Infant and Child Development, n/a*(n/a), Article e2370. https://doi.org/10.1002/icd.2370

von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing, 17*(4), 395–416. https://doi.org/10.1007/s11222-007-9033-z

Wang, B., Mezlini, A. M., Demir, F., Fiume, M., Tu, Z., Brudno, M., … Goldenberg, A. (2014). Similarity network fusion for aggregating data types on a genomic scale. *Nature Methods, 11*(3), 333–337. https://doi.org/10.1038/nmeth.2810

Wang, H.-T., Smallwood, J., Mourao-Miranda, J., Xia, C. H., Satterthwaite, T. D., Bassett, D. S., & Bzdok, D. (2020). Finding the needle in a high-dimensional haystack: Canonical correlation analysis for neuroscientists. *NeuroImage, 216*, Article 116745. https://doi.org/10.1016/j.neuroimage.2020.116745

Winkler, A. M., Renaud, O., Smith, S. M., & Nichols, T. E. (2020). Permutation inference for canonical correlation analysis. *NeuroImage, 220*, Article 117065. https://doi.org/10.1016/j.neuroimage.2020.117065

Witten, D. M., Tibshirani, R., & Hastie, T. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics (Oxford, England), 10*(3), 515–534. https://doi.org/10.1093/biostatistics/kxp008

Xu, D., & Tian, Y. (2015). A comprehensive survey of clustering algorithms. *Annals of Data Science, 2*(2), 165–193. https://doi.org/10.1007/s40745-015-0040-1

Ying, X. (2019). An overview of overfitting and its solutions. *Journal of Physics: Conference Series, 1168*(2), Article 022022. https://doi.org/10.1088/1742-6596/1168/2/022022

Yu, Q., Risk, B. B., Zhang, K., & Marron, J. S. (2017). JIVE integration of imaging and behavioral data. *NeuroImage, 152*, 38–49. https://doi.org/10.1016/j.neuroimage.2017.02.072

Yu, & Shi. (2003). Multiclass spectral clustering. In *, Vol. 1*. *Proceedings ninth IEEE international conference on computer vision* (pp. 313–319). https://doi.org/10.1109/ICCV.2003.1238361

Zhuang, X., Yang, Z., & Cordes, D. (2020). A technical review of canonical correlation analysis for neuroscience applications. *Human Brain Mapping, 41*(13), 3807–3833. https://doi.org/10.1002/hbm.25090

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 67*(2), 301–320.