



Enhancing Within-Person Estimation of Neurocognition and the Prediction of Externalizing Behaviors in Adolescents

RESEARCH ARTICLE

SAM PASKEWITZ 

INTI A. BRAZIL 

ILKER YILDIRIM 

SONIA RUIZ 

ARIELLE BASKIN-SOMMERS 

*Author affiliations can be found in the back matter of this article

][ubiquity press

ABSTRACT

Decades of research document an association between neurocognitive dysfunction and externalizing behaviors, including rule-breaking, aggression, and impulsivity. However, there has been very little work that examines how multiple neurocognitive functions co-occur within individuals and which combinations of neurocognitive functions are most relevant for externalizing behaviors. Moreover, Latent Profile Analysis (LPA), a widely used method for grouping individuals in person-centered analysis, often struggles to balance the tradeoff between good model fit (splitting participants into many latent profiles) and model interpretability (using only a few, highly distinct latent profiles). To address these problems, we implemented a non-parametric Bayesian form of LPA based on the Dirichlet process mixture model (DPM-LPA) and used it to study the relationship between neurocognitive functioning and externalizing behaviors in adolescents participating in the Adolescent Brain Cognitive Development Study. First, we found that DPM-LPA outperformed conventional LPA, revealing more distinct profiles and classifying participants with higher certainty. Second, latent profiles extracted from DPM-LPA were differentially related to externalizing behaviors: profiles with deficits in working memory, inhibition, and/or language abilities were robustly related to different expressions of externalizing. Together, these findings represent a step towards addressing the challenge of finding novel ways to use neurocognitive data to better describe the individual. By precisely identifying and specifying the variation in neurocognitive and behavioral patterns this work offers an innovative empirical foundation for the development of assessments and interventions that address these costly behaviors.

CORRESPONDING AUTHOR:

Sam Paskewitz

Department of Psychology, Yale University, New Haven CT, US

samuel.paskewitz@yale.edu

KEYWORDS:

neurocognition; Bayesian latent profile analysis; externalizing behaviors; adolescents

TO CITE THIS ARTICLE:

Paskewitz, S., Brazil, I. A., Yildirim, I., Ruiz, S., & Baskin-Sommers, A. (2024). Enhancing Within-Person Estimation of Neurocognition and the Prediction of Externalizing Behaviors in Adolescents. *Computational Psychiatry*, 8(1), pp. 119–141. DOI: <https://doi.org/10.5334/cpsy.112>

Externalizing behaviors consist of aggressive, rule-breaking, destructive, and deceitful behaviors and are expressed in approximately a third of school-aged youth (Danielson et al., 2021). For many children, externalizing behaviors begin during preschool or early school years. However, children are most at risk for externalizing behaviors during adolescence, a developmental period characterized by significant biological and psychological changes that coincide with increased social challenges (Beauchaine & McNulty, 2013; Moffitt, 1993). Adolescents displaying externalizing behaviors are at increased risk for academic underachievement, family dysfunction, legal system involvement, substance misuse, emotional distress, suicidality, teen pregnancy, and a host of health problems (Bongers, Koot, van der Ende, & Verhulst, 2008; Timmermans, van Lier, & Koot, 2008; Van der Ende, Verhulst, & Tiemeier, 2016). Consequently, externalizing behaviors are associated with a high individual, family, and societal burden; they constitute a leading reason for referral to mental health services among adolescents and are a main cause of disability worldwide (Shepherd & Farrington, 2003). Given the profound burden of externalizing behaviors, determining the factors that underlie these behaviors is an essential step for advancing prevention and treatment development.

NEUROCOGNITION AND EXTERNALIZING BEHAVIORS

A substantial body of research indicates that neurocognitive differences underlie the onset and maintenance of externalizing behaviors. Neurocognitive models of externalizing behaviors most commonly highlight the relevance of various components of executive functions, a term that encompasses cognitive processes related to the initiation planning, and regulation of behavior (Honrath et al., 2021). Meta-analyses of cross-sectional (Ogilvie, Stewart, Chan, & Shum, 2011) and longitudinal (Yang et al., 2022) data report medium-sized effects showing deficits in executive functioning, particularly within subcomponents of working memory and inhibition. However, neurocognition is a multicomponent construct and other components, such as learning, memory, and general decision-making, also appear deficient among adolescents showing externalizing behaviors (Fairchild et al., 2019; Van Goozen, Langley, & Hobson, 2022). For example, Thompson and colleagues (2019) performed factor analysis on neurocognitive data from a large sample of 9–10-year-olds and found that low scores on the general cognitive ability and learning/memory functions, not just executive functioning, were associated with more externalizing behaviors.

Nonetheless, many studies in this area only include one or two subcomponents of neurocognition and few studies explore components of neurocognition beyond executive functioning. Further, much of the extant research relies on analytic methods (multiple regression, factor analysis) that struggle to capture complex, within-person, interactions among components of neurocognition, despite evidence that neurocognitive components work in concert within each person (Chaku, Barry, Fowle, & Hoyt, 2022; Ruiz, Brazil, & Baskin-Sommers, 2023). Part of the challenge lies in finding effective analytical and conceptual frameworks to better describe variability in neurocognition, both within and across individuals (Brazil, van Dongen, Maes, Mars, & Baskin-Sommers, 2018).

Latent profile analysis (LPA) is a popular method for studying within-person variation. LPA explains observed variables by grouping participants into latent profiles, i.e. categories of people with similar characteristics. In contrast to factor analysis, LPA does not assume homogenous relationships among variables: it can represent groups of participants who perform neurocognitive tasks in different ways. For example, a recent study by Chaku and colleagues (2022) applied LPA to three neurocognitive variables representing different aspects of executive functioning (flanker task for inhibition, list sorting for working memory, and card sorting for cognitive flexibility). They identified four profiles, representing groups with overall high, medium, and low executive function, respectively, as well as one with specifically impaired inhibition. Adolescents belonging to the low executive functioning profile showed more externalizing behaviors than others.

Despite its advantages for estimating within-person neurocognitive functioning, conventional LPA has a significant limitation: choosing the number of latent profiles involves a tradeoff between model fit and model interpretability (Figure 1 displays a schematic representation). An LPA model with too many latent profiles (Figure 1b) will fit the data well (i.e. have a larger log-likelihood) by producing

a fine-grained description of participants. However, this comes at the cost of interpretability: some profiles will be very similar. By comparison, an LPA model with too few latent profiles (Figure 1c) will have distinctive profiles and thus be easy to interpret. However, some participants will not be well described by any of the profiles, resulting in poorer model fit (log-likelihood).

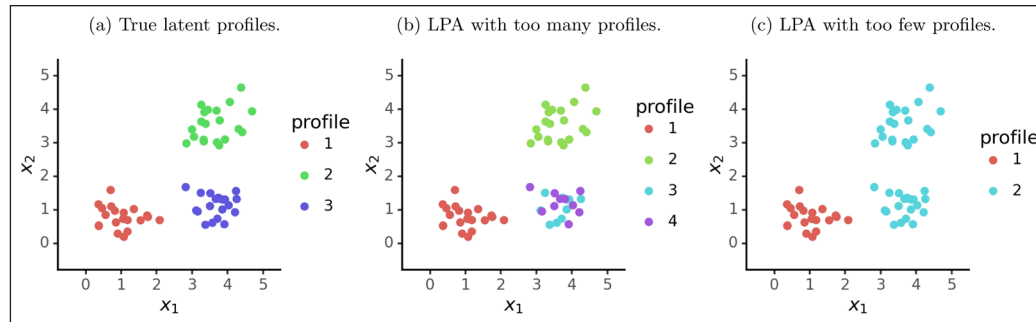


Figure 1 Illustration of LPA with too few or too many profiles.

To determine the correct number of latent profiles, researchers fit multiple LPA models, each with a different number of profiles, and compare them using criteria such as model fit, classification uncertainty, a minimum number of people in each profile, statistical tests, and profile interpretability. However, these criteria often disagree, leaving model selection up to the researcher's judgment (e.g. Chaku et al., 2022; Wendt et al., 2019). The number of profiles selected dramatically influences the interpretation of an LPA model and, by extension, how we understand the relationship between neurocognition and externalizing.

USING NON-PARAMETRIC BAYESIAN INFERENCE TO BALANCE THE LPA FIT-INTERPRETABILITY TRADEOFF

In the present study, we sought to develop a non-parametric Bayesian LPA model to address limitations of the conventional LPA model and estimate within-person neurocognitive functioning. The Bayesian framework offers a solution by providing flexible probabilistic modeling approaches that rely less on assumptions and conventions and instead, allow us to use data-derived probability distributions to generate inferences (see Van de Schoot et al., 2021, for an introduction).

Bayesian inference uses both the model likelihood and a prior distribution to estimate a model. The prior distribution describes the inherent probability of different model solutions. For LPA, Bayesian inference requires a prior distribution of participants' latent profile membership. Existing Bayesian LPA methods use a prior that assumes a fixed number of latent profiles (Li, Lord-Bessen, Shiyko, & Loeb, 2018; White & Murphy, 2014). They thus suffer from a similar drawback to conventional LPA: to decide the correct number of latent profiles, one must fit and compare models of different sizes. We avoid this problem by using a distribution called the Dirichlet process (Ferguson, 1973) as the prior on profile membership. The resulting model (a Dirichlet process mixture) is very flexible: it allows for a large number of latent profiles, but places higher prior probability on model fits with fewer profiles. We developed a novel implementation of the Dirichlet mixture for application to latent profile analysis (DPM-LPA). DPM-LPA has the flexibility to infer a large number of latent profiles if needed, but unlike conventional LPA, DPM-LPA favors simpler model fits that produce a small number of distinct, non-redundant profiles (see Method; see also Qiu, Paganin, Ohn, & Lin, 2023, for a computationally different implementation of the same idea).

Here, the first goal was to compare DPM-LPA and conventional LPA. We fit both models to neurocognitive data from 11–12-year-olds included in the Adolescent Brain Cognitive Development StudySM (ABCD Study[®]). Every two years participants complete the NIH Toolbox task battery (Luciana et al., 2018) and several other neurocognitive tasks during a magnetic resonance imaging session (Casey et al., 2018). We compared DPM-LPA and conventional LPA with respect to two metrics: profile similarity and the certainty with which participants were classified into profiles (entropy reduction). We hypothesized that DPM-LPA would perform better on both metrics, producing more interpretable profiles than conventional LPA. In addition, we compared DPM-LPA to conventional LPA and finite Bayesian LPA in a simulation study to determine how accurately these different

methods could infer the true number of latent profiles. We hypothesized that DPM-LPA would perform at least as well as the other methods.

The second goal was to validate the DPM-LPA neurocognitive latent profiles in relation to externalizing behaviors. Participants complete measures of mental health symptoms and behaviors (Barch et al., 2018), including the Child Behavior Checklist (CBCL) measure of externalizing behaviors (Achenbach, 2009) and the Positive and Negative Urgency Scales (Cyders et al., 2007; Whiteside, Lynam, Miller, & Reynolds, 2005), which measure affective-based impulsive behavior. We used Bayesian methods to investigate the relationship among the neurocognitive profiles discovered by DPM-LPA and externalizing behaviors. We expected that profiles characterized by worse neurocognitive performance would have higher average levels of externalizing but given that we present a novel application of DPM-LPA, we did not have more specific hypotheses about the relationships among neurocognitive profiles and externalizing behaviors.

METHOD

PARTICIPANTS

Participants were youth included in the ABCD Study Data Release 5.0 with a complete set of neurocognitive variables at the 2-year follow-up (T2; ages 11–12) ($n = 5232$; 47.9% Female, 52.1% Male; 1.9% Asian; 11.0% Black; 18.9% Hispanic; 10.2% Other; 57.9% White). We compared participants with a complete set of neurocognitive variables to those with missing neurocognitive data with respect to sex, race/ethnicity, externalizing behaviors, and internalizing behaviors (see Supplemental Material for details). The only difference was with respect to race/ethnicity: the proportion of Black and Hispanic youth was lower among participants with complete neurocognitive data (Black: 11.0% vs. 18.2%, Hispanic: 18.9% vs. 21.4%). All parents or caregivers provided written informed consent and children provided verbal assent for participation in the study (Clark et al., 2018). See Garavan and colleagues (2018) for the ABCD Study baseline exclusion criteria.

ASSESSMENTS

The ABCD Study data collection involves biannual visits an extensive evaluation across neurocognition and behavior (Jernigan, Brown, & Dowling, 2018). We used data from the 2-year and 3-year follow-up visits across all 21 ABCD sites (data release 5.0, doi: [10.15154/8873-zj65](https://nda.nih.gov/study.html?id=901); <https://nda.nih.gov/study.html?id=901>). Figure 2 provides correlations among all variables.

Neurocognitive Data

We selected 11 neurocognitive measures (Table 1a) that spanned different domains of neurocognition, including executive functions, learning/memory, and general cognition (Thompson et al., 2019). For all NIH toolbox tasks, we used the uncorrected standardized scores to summarize behavioral performance given that there has been some question about the use of the T-scores for neurocognitive assessment in adolescents (Luciana et al., 2018; Taylor et al., 2022).

Executive functions

From the NIH Toolbox, we included Flanker (cognitive/attentional control) and Pattern Completion (processing speed). We also included a measure from the Stop Signal Task, which assesses response inhibition (Logan, 1994). On most trials, the participant's objective is to quickly give one of two responses depending on the direction of an arrow. However, a subset of trials feature a stop signal, indicating that the response should be withheld. Analysis of response accuracy and response times allows computation of a stop signal reaction time (SSRT) that measures how well each participant can inhibit their responses; we used this measure (reverse coded by multiplying it by -1). Lastly, we included accuracy from the Emotional N-Back Test. The task includes 0-back blocks (respond to a stimulus seen at the beginning of the block) and 2-back blocks (respond to the stimulus seen two trials ago). The task includes both emotionally neutral stimuli (buildings) and emotionally charged stimuli (faces). We used accuracy across stimuli type during 2-back trials as a measure of working memory.

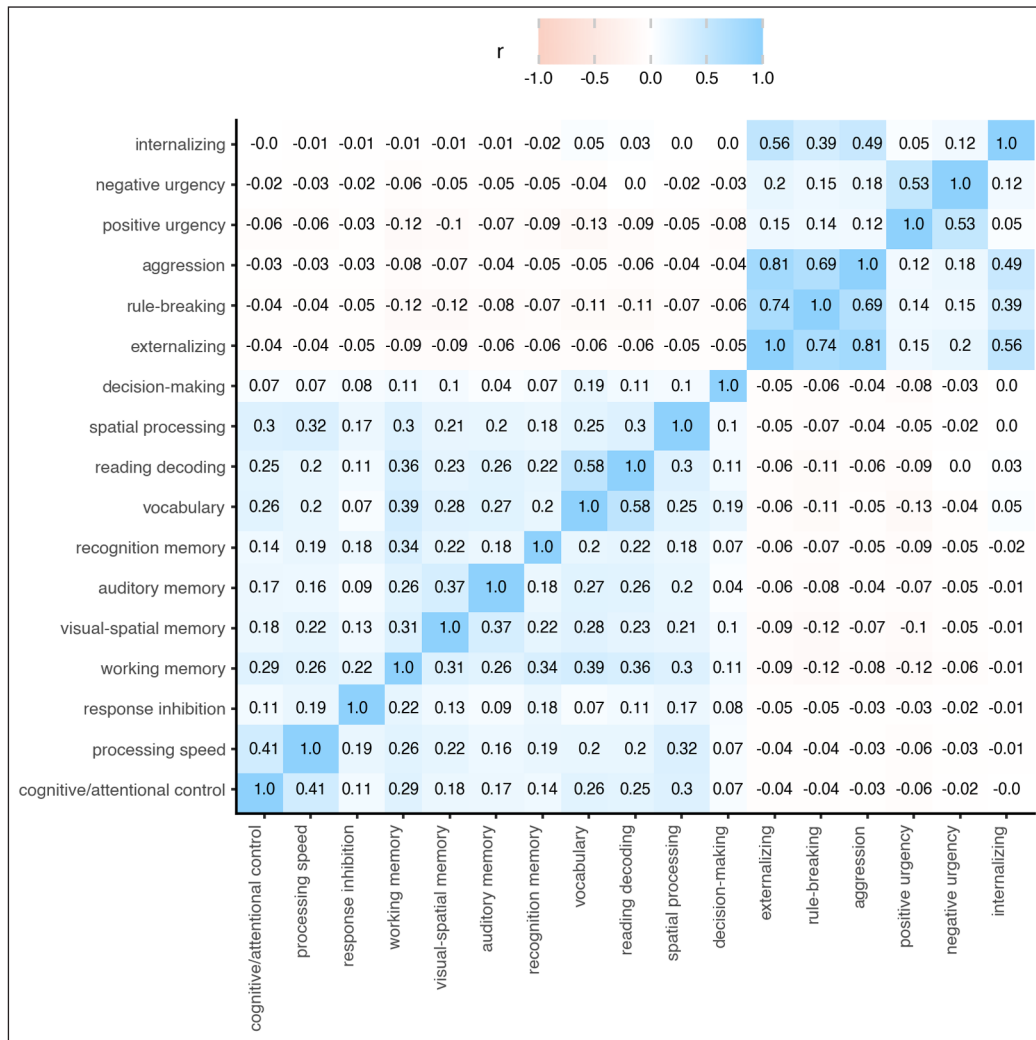


Figure 2 Correlations between variables (two year follow-up).

Table 1 Variables.

(A) NEUROCOGNITIVE (INDICATOR) VARIABLES (x)		
COGNITIVE PROCESS	TASK	DESCRIPTION
Executive Function		
cognitive/attentional control	Flanker (NIH Toolbox)	uncorrected standard score
processing speed	Pattern Comparison (NIH Toolbox)	uncorrected standard score
response inhibition	Stop Signal	-1× stop signal reaction time
working memory	Emotional N-Back	proportion correct on two-back trials
Learning and Memory		
visual-spatial memory	Picture Sequence Memory (NIH Toolbox)	uncorrected standard score
auditory memory	Rey Auditory Verbal Learning	immediate recall total correct
recognition memory	Emotional N-Back	recognition trials d'
General Cognition		
vocabulary	Picture Vocabulary (NIH Toolbox)	uncorrected standard score
reading decoding	Oral Reading Recognition (NIH Toolbox)	uncorrected standard score
spatial processing	Little Man Task	% correct/avg. correct response time
decision-making	Game of Dice	# safe choices - # risky choices

(Contd.)

(B) OUTCOME VARIABLES (y)

VARIABLE	SAMPLE SIZE (2-YEAR FOLLOW-UP)	SAMPLE SIZE (3-YEAR FOLLOW-UP)
externalizing (CBCL)	4977	4742
rule-breaking (CBCL)	4977	4742
aggression (CBCL)	4977	4742
internalizing (CBCL)	4977	4742
positive urgency (UPPS-P)	5230	–
negative urgency (UPPS-P)	5230	–

Learning and Memory

From the NIH Toolbox, we included Picture Sequence Memory (visuospatial sequencing/memory). To measure auditory memory, we used the number of correct words in the immediate recall test from the Rey Auditory Verbal Learning Task (Strauss, Sherman, & Spreen, 2006). Finally, in a second phase of the Emotional N-Back Test, a test of recognition memory was administered where participants are asked to discriminate between stimuli seen in the previous stage and new ones. We used d' , a discriminability metric from signal detection theory (Green & Swets, 1966) to measure recognition memory.

General Cognition

From the NIH Toolbox, we included Picture Vocabulary (measuring language skills/verbal intellect) and Oral Reading Recognition (language skills/reading decoding). We also used the Little Man Task (Acker & Acker, 1982) which measures visual-spatial processing. On each trial, participants see a rudimentary male figure rotated in various positions and must determine in which hand the figure is holding a briefcase. Performance is summarized by an efficiency score computed as percent correct responses divided by average response time. Lastly, to assess decision-making, we included the Game of Dice Task (Brand et al., 2005), which was adapted from the Iowa Gambling Task. On each trial, the participant chooses between a low-risk/low reward gamble and a high-risk/high-reward gamble. High-risk gambles are designed to be disadvantageous, so decision-making performance is measured by the number of low risk choices minus the number of high risk choices.

This set of neurocognitive measures differed from those examined in previous work using the ABCD data (Chaku et al., 2022; Thompson et al., 2019). We analyzed data from when participants were 11–12 years old, whereas similar previous work looked at baseline data (9–10 years old). Over that period, the set of neurocognitive tasks administered to ABCD participants changed: the Card Sort (cognitive flexibility) and List Sort (working memory) tasks were dropped after the baseline assessment, and the Game of Dice Task (decision-making) was added starting at the two-year-follow-up timepoint. In addition, we used behavioral data from the Stop Signal and Emotional N-Back tasks. All neurocognitive variables were standardized by subtracting the mean and dividing by the standard deviation (i.e. z-scored) prior to analysis.

Outcome Variables

We used the parent-report CBCL (Achenbach, 2009) and the UPPS (Cyders et al., 2007; Whiteside, Lynam, Miller, & Reynolds, 2005) to measure externalizing behaviors (Table 1b). The CBCL provides a total externalizing score, and two subscales representing rule-breaking and aggressive behaviors. The CBCL was collected at the two-year-follow-up timepoint (concurrent with the neurocognitive data) and at the three-year-follow-up timepoint (one year after the neurocognitive data). We also performed a sensitivity analysis including the broad internalizing measure from the CBCL (Achenbach, 2009), which measures anxiety, depression/withdrawal, and somatic complaints. We opted to use the parent-report for these measures because the youth-report assessment for the ABCD Study protocol changed across the timepoints (Barch et al., 2018), especially in the externalizing measures. In addition, research has documented

good concordance between youth- and parent-report measures of externalizing disorders, which tend to be observable behaviors (note: concordance for internalizing tends to be lower than for externalizing in this age group; [De Los Reyes & Kazdin, 2005](#)). The UPPS scale provides two measures of impulsivity in affective circumstances (UPPS-Negative Urgency: the tendency to act rashly when experiencing negative emotions; UPPS-Positive Urgency: tendency to act rashly when experiencing positive emotion). The UPPS was only collected at the two-year-follow-up timepoint. All outcome variables were standardized by subtracting the mean and dividing by the standard deviation (i.e. z-scored).

LATENT PROFILE ANALYSIS

Conventional Latent Profile Analysis (LPA)

We used the tidyLPA and mclust packages in R to perform conventional LPA ([Rosenberg, Beymer, Anderson, Van Lissa, & Schmidt, 2019](#); [Scrucca, Fraley, Murphy, & Raftery, 2023](#)). We fit models with numbers of profiles ranging from 1 to 20. While it is rare to fit a conventional LPA model with 20 profiles, DPM-LPA can infer a large number of latent profiles if needed to describe the data, so including large conventional LPA models made for a fairer comparison. In mathematical terms, LPA is a finite mixture of multivariate Gaussians with a shared diagonal covariance matrix. The model starts by assuming that there is a fixed number of latent profiles (T) and that each individual belongs to a single latent profile. We use z_i to represent person i 's latent profile. For example, if person i belongs to latent profile 3, then $z_i = 3$. Latent profiles can be arbitrarily relabeled, so it is convention to label the profile containing the largest number of participants as profile 1, the next largest profile as profile 2, etc. Each person has a set of m observed indicator variables that we denote $x_i = x_{1,i}, x_{2,i}, \dots, x_{m,i}$. In our study, these indicator variables are performance measures on the various neurocognitive tasks described in [Table 1a](#). Each person's indicator variables (x_i) depend on the latent profile that person belongs to (z_i). In particular, each indicator variable ($x_{j,i}$) is assumed to have a normal distribution with a mean ($\mu_{j,t}$) that varies based on latent profile:

$$x_{j,i}|z_i = t \sim \mathcal{N}(\mu_{j,t}, \frac{1}{\xi_j}) \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, m \quad (1)$$

Each profile is thus defined by its vector of means ($\mu_t = \mu_{1,t}, \mu_{2,t}, \dots, \mu_{m,t}$). Because all of our indicator variables are standardized (i.e. they have mean 0 and variance 1), it means that if $\mu_{j,t} > 0$ then people in profile t tend to have above average values of indicator variable $x_{j,p}$ while if $\mu_{j,t} < 0$ they have below average values.

ξ_j represents the precision, i.e. inverse variance, of indicator variable j : $\text{var}(x_{j,i}) = \frac{1}{\xi_j}$. Thus, a variable with high precision has low variance, and vice versa (parameterization in terms of precision makes computations more convenient). We assume that precision (variance) does not differ across latent profiles. This is largely a pragmatic assumption: if precision did vary across profiles, it would be difficult to estimate it for smaller profiles (those with fewer people). We also assume that the indicator variables ($x_{j,i}$) do not have any covariance with each other, i.e., they are independent given latent profile membership (z_i). This is important for making latent profile analysis easy to interpret: one only needs to know a participant's latent profile to know the distribution of any of their indicator variables ($x_{j,i}$). The model has the following parameters that must be inferred:

$$\xi_j = \text{precision (inverse variance) of indicator } j \quad j = 1, 2, \dots, m \quad (2)$$

$$\mu_{j,t} = \text{mean of indicator } j \text{ in profile } t \quad j = 1, 2, \dots, m, \quad t = 1, 2, \dots, T \quad (3)$$

$$\pi_t = \text{base rate probability of profile } t \quad t = 1, 2, \dots, T \quad (4)$$

This conventional LPA model is estimated using the expectation maximization (EM) algorithm ([Dempster, Laird, & Rubin, 1977](#)) to estimate these parameters by maximizing the model likelihood. This also produces estimates of z (profile membership) in the form of probability vectors (ϕ_i):

$$\phi_i = [\phi_{1,i}, \phi_{2,i}, \dots, \phi_{T,i}] = \text{probabilities that participant } i \text{ is in each profile} \quad (5)$$

The conventional LPA model assumes a fixed number of latent profiles (T). Thus to determine how many latent profiles are the data one must fit multiple models with different values of T and determine which best describes the data. We examine three standard criteria for comparing models to choose the correct number of profiles. First, the entropy reduction statistic quantifies how confident each model is about its classification of people into different latent profiles. We describe it in more detail below. A low entropy reduction statistic (less than about 0.8) suggests that the model cannot accurately assign participants to latent profiles (Wang, Deng, Bi, Ye, & Yang, 2017). Second, the Akaike Information Criterion (AIC, Akaike, 1974) and Bayesian Information Criterion (BIC, Schwarz, 1978) measure goodness of fit (log likelihood) penalized by model complexity (number of estimated parameters). Lower values indicate better model fit balanced by model complexity. Finally, the Bootstrap Likelihood Ratio Test (Langeheine, Pannekoek, & Van de Pol, 1996) provides a test of whether – for each number of profiles – adding an additional profile improves model fit.

Dirichlet Process Mixture Latent Profile Analysis (DPM-LPA)

DPM-LPA is a non-parametric Bayesian form of LPA in which the number of latent profiles is not specified beforehand. Conventional LPA finds point estimates of model parameters by maximizing the likelihood of the data. In contrast, Bayesian models are fit by computing a probability distribution over their parameters (including participant-specific variables such as latent profile membership, z_i) rather than a single point estimate. This distribution, the posterior distribution, combines a prior distribution (representing what parameter values are probable before data are observed) and the likelihood via Bayes' Rule:

$$\overbrace{p(z_{1:n}, \mu_{1:T}, \xi_{1:m} | x_{1:n})}^{\text{posterior}} = \frac{\overbrace{p(x_{1:n} | z_{1:n}, \mu_{1:T}, \xi_{1:m})}^{\text{likelihood}} \overbrace{p(z_{1:n}, \mu_{1:T}, \xi_{1:m})}^{\text{prior}}}{\underbrace{p(x_{1:n})}_{\text{evidence}}} \quad (6)$$

Thus, in Bayesian statistics, parameter estimates are influenced by both the likelihood (as in conventional LPA) and the prior distribution. The prior distribution often leads to simpler model fits than would be obtained by maximum likelihood estimation (Jefferys & Berger, 1992).

DPM-LPA uses the same likelihood function as conventional LPA (Equation 1): each person's vector of observed indicator variables (x_i) follows a normal distribution with a mean vector (μ) that depends on the latent profile that person belongs to (z_i) and with precisions (ξ) that are shared across profiles.

The prior distribution for DPM-LPA can be broken into two parts, the prior over latent profile membership, and the prior over means/precisions:

$$\overbrace{p(z_{1:n}, \mu_{1:T}, \xi_{1:m})}^{\text{full prior}} = \overbrace{p(z_{1:n})}^{\text{profile membership}} \overbrace{p(\mu_{1:T}, \xi_{1:m} | z_{1:n})}^{\text{means/precisions}} \quad (7)$$

The key to DPM-LPA is its prior on profile membership (z_i), which is a distribution called the Dirichlet process (Ferguson, 1973). Using the Pólya Urn scheme representation (Blackwell & MacQueen, 1973), we can write it as:

$$p(z_{1:n}) = \frac{\alpha^T \prod_{t=1}^T (n_t - 1)!}{\alpha(\alpha + 1) \cdots (n - 1 + \alpha)} \quad n_t = \text{number of participants in profile } t \quad (8)$$

where α is a concentration parameter (we infer α from the data, following Blei & Jordan, 2006). This Dirichlet process prior distribution does not have a hard upper limit on the number of latent profiles. However, it favors simpler LPA solutions. The prior on $z_{1:n}$ is high when participants are concentrated in a few latent profiles and low when they are spread across several large ones, with this preference varying as a function of α (Figure 3). A more complex model fit, with participants assigned to a greater number of latent profiles, will only be favored (have a higher posterior probability) if the resulting increase in likelihood by adding those profiles outweighs the decrease in the prior distribution.

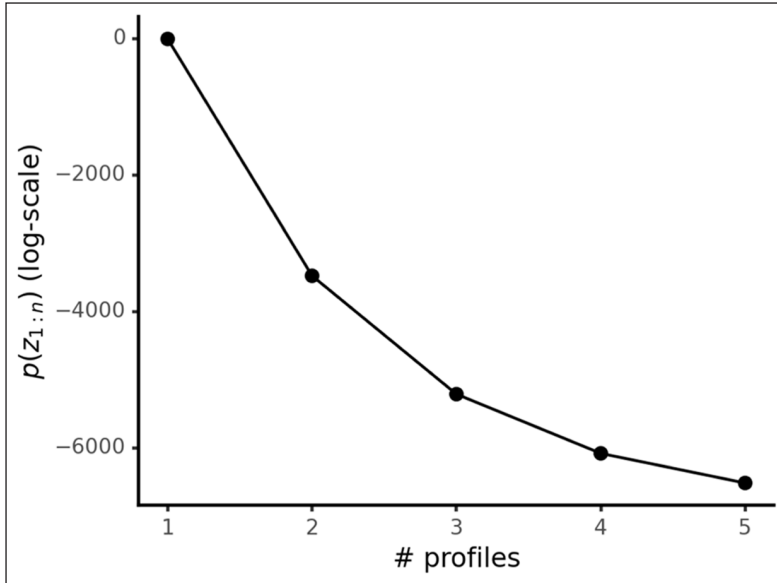


Figure 3 DPM-LPA prior distribution (log-scale) on participant latent profile membership ($z_{1:n}$) as a function of the number of latent profiles. Calculated with $n = 5000$ participants, with each new profile created by splitting the last one in half.

Means/precisions have a standard form of prior (called a conjugate prior) that keeps computations simple:

$$p(\mu_{1:T}, \xi_{1:m} | z_{1:n}) = \prod_{j=1}^m p(\xi_j) \prod_{t=1}^T p(\mu_{j,t} | \xi_j) \quad (9)$$

$$\xi_j \sim \text{Gamma}(5, 5) \quad j = 1, 2, \dots, m \quad (10)$$

$$\mu_{j,t} | \xi_j \sim \mathcal{N}(0, \frac{2}{\xi_j}) \quad j = 1, 2, \dots, m \quad t = 1, 2, \dots, T \quad (11)$$

where T represents the number of distinct latent profiles implied by $z_{1:n}$.

DPM-LPA is complex enough that its posterior distribution must be approximated. Markov chain Monte Carlo is one common approximation technique that produces a set of random samples from the posterior distribution (Ishwaran & James, 2001; Qiu et al., 2023). However, this is computationally expensive and thus not attractive for large datasets. Instead, we used an approximation method called mean field variational Bayes (for a more thorough discussion see Blei & Jordan, 2006; Blei, Kucukelbir, & McAuliffe, 2017). We implemented this algorithm and supporting code in Python. Once the DPM-LPA model has been fit, it is easy to determine how many latent profiles describe the data. Most of the latent profiles estimated by the model will be empty, i.e. the total estimated probability of any participants belonging to them will be approximately zero. We used the simple rule that if a profile was not estimated as maximally probable for at least one participant, that profile was considered to be empty and excluded from further analysis and discussion.

Entropy Reduction Statistic

We used a standard entropy reduction statistic to compare how confidently the conventional LPA and DPM-LPA models classify people into latent profiles. For each participant, the model estimates the probability of that person belonging to each latent profile: we call this ϕ_i . For example, if for a certain participant $\phi_i = [0.9, 0.07, 0.03]$ then the model estimates that the probability of this person belonging to profile 1 is 90%, to profile 2 is 7%, and to profile 3 is 3% (in this example there are three latent profiles). Intuitively, the model is fairly confident that the participant is in profile 1. If the probabilities were closer to being equal, say for example $\phi_i = [0.4, 0.32, 0.28]$, then the model would be less confident about which profile the participant belonged to. The entropy of ϕ_i quantifies this uncertainty:

$$\text{entropy for participant } i = - \sum_{t=1}^T \phi_{t,i} \log(\phi_{t,i}) \quad (12)$$

Consider the entropy of three hypothetical participants. In the first case ($\phi_1 = [0.9, 0.07, 0.03]$) entropy is low (entropy = 0.39), meaning that the model is fairly certain. In the second case ($\phi_2 = [0.4, 0.32, 0.28]$) entropy is high (entropy = 1.09), meaning that the model is not very certain. Maximum entropy occurs when all the values in ϕ_i are equal, meaning intuitively that the model does not have any idea which profile the participant belongs to ($\phi_3 = [0.33, 0.33, 0.33]$, entropy = 1.10 in a model with three profiles).

The proportional entropy reduction statistic quantifies the overall certainty of a model. It compares the total entropy of a model across participants ($-\sum_{i=1}^n \sum_{t=1}^T \phi_{t,i} \log(\phi_{t,i})$) to the maximum possible entropy ($-n \log(\frac{1}{T})$). Subtracting this ratio from 1 gives the following statistic:

$$\text{entropy reduction} = \frac{\text{max entropy} - \text{total entropy}}{\text{max entropy}} = 1 - \frac{\text{total entropy}}{\text{max entropy}} = 1 - \frac{\sum_{i=1}^n \sum_{t=1}^T \phi_{t,i} \log(\phi_{t,i})}{n \log(\frac{1}{T})} \quad (13)$$

Values can range from 1 (total confidence in classifying participants) down to 0 (the opposite).

Profile Distinctiveness (Mahalanobis Distance)

Ideally, the latent profiles discovered by LPA should be very distinct from one another. This makes the profiles easier to interpret and reduces uncertainty about participants' profile membership. We used the Mahalanobis distance to quantify the distinctiveness between each pair of profiles. This is a metric based on the Euclidean distance between two profiles' mean vectors (μ_1 and μ_2), scaled by the precision, i.e. inverse variance (ξ):

$$\text{distance} = \sqrt{\sum_{j=1}^m \xi_j (\mu_{j,1} - \mu_{j,2})^2} \quad (14)$$

A distance of 0 indicates that two profiles are identical, while distance increases as profiles become more and more distinct. For each model (conventional LPA of varying sizes and DPM-LPA) we computed the minimum distance (to find the two most similar profiles in each model) as well as the mean distance across profile pairs.

SIMULATIONS

We compared how well DPM-LPA, finite Bayesian LPA, and conventional LPA could infer the true number of latent profiles in simulated data. The simulation parameters were a subset of those used in a study by Tein, Cox, and Cham (2013) examining various criteria for determining the number of profiles in conventional LPA. See the Supplemental Material for further details.

ANALYZING THE RELATIONSHIP BETWEEN LATENT PROFILES AND OUTCOME VARIABLES

Estimating Profile Means ($\mu^{(y)}$)

We used Bayesian methods to analyze the relationship between DPM-LPA latent profiles and outcome variables. For the sake of clarity, let us refer to the set of indicator variables (in our case neurocognitive measures) as x and the outcome as y . We assume that outcomes (e.g., CBCL externalizing) follow a normal distribution with a mean ($\mu^{(y)}$) that depends on the latent profile but a shared precision (ξ_y):

$$y_i | z_i = t \sim \mathcal{N}(\mu_t^{(y)}, \frac{1}{\xi^{(y)}}) \quad i = 1, 2, \dots, n \quad (15)$$

We used a normal-gamma conjugate prior distribution:

$$\xi^{(y)} \sim \text{Gamma}(1, 1) \quad (16)$$

$$\mu_t^{(y)} | \xi^{(y)} \sim \mathcal{N}(0, \frac{1}{\xi^{(y)}}) \quad t = 1, 2, \dots, T \quad (17)$$

If each participant's true latent profile (z_i) was known, then the formula for estimating the outcome means ($\mu^{(y)}$) for each latent profiles would be standard conjugate prior updates.

However, we of course do not know z_i , but only have an estimate from the DPM-LPA fit in the form of a probability vector (ϕ_i). We followed a simple, commonly used procedure (e.g. Chaku et al., 2022) and assigned participants to their most probable profiles:

$$\text{estimated profile for participant } i = \hat{z}_i = \underset{t=1,2,\dots,T}{\operatorname{argmax}}(\phi_{t,i}) \quad i = 1, 2, \dots, n \quad (18)$$

We then use these profile estimates ($\hat{z}_i$) in the ordinary formulas to obtained approximate posterior distributions over each outcome's variance and profile means. This also gave us 95% posterior credible intervals over these parameters (the interval between the 2.5th and 97.5th percentiles of the posterior distribution).

Comparing Profile Means

We used a Bayesian analogue of the analysis of variance (ANOVA) to test if any of the profiles differed from each other, and if they did, we followed up with Bayesian post-hoc tests to determine which profiles had different means. The Bayesian ANOVA (Rouder, Morey, Speckman, & Province, 2012) compares a null hypothesis H_0 (which assumes that all profiles have equal means) to the alternative hypothesis H_1 that profile means differ:

$$H_0 : \mu_1^{(y)} = \mu_2^{(y)} = \dots = \mu_T^{(y)} \quad (19)$$

$$H_1 : \mu_1^{(y)}, \mu_2^{(y)}, \dots, \mu_T^{(y)} \text{ are not all equal} \quad (20)$$

We compared these two hypotheses using the Bayes factor (BF_{10}), which is defined as the ratio between the evidence for the two hypotheses:

$$BF_{10} = \frac{p(y_{1:n}|H_1)}{p(y_{1:n}|H_0)} \quad (21)$$

Where the evidence for either hypothesis is computed as the likelihood of the observed data ($y_1, y_2, \dots, y_n$) averaged across the prior distribution:

$$p(y_{1:n}|H_0) = \int_{\xi^{(y)}} \int_{\mu^{(y)}} p(y_{1:n}|\xi^{(y)}, \mu^{(y)}) p(\xi^{(y)}) p(\mu^{(y)}) d\mu^{(y)} d\xi^{(y)} \quad (22)$$

$$p(y_{1:n}|H_1) = \int_{\xi^{(y)}} \int_{\mu_1^{(y)}} \dots \int_{\mu_T^{(y)}} p(y_{1:n}|\xi^{(y)}, \mu_1^{(y)}, \dots, \mu_T^{(y)}) p(\xi^{(y)}) p(\mu_1^{(y)}) \dots p(\mu_T^{(y)}) d\mu_T^{(y)} \dots d\mu_1^{(y)} d\xi^{(y)} \quad (23)$$

Bayes factors greater than 1 show support for H_1 , Bayes factors less than 1 show support for H_0 , and Bayes factors close to 1 show similar amounts of evidence for both hypotheses.

Because Bayes factors can vary so widely in size, we also reported them in base-10 logarithmic scale:

$$\log_{10}(BF_{10}) = \log_{10}(p(y_{1:n}|H_1)) - \log_{10}(p(y_{1:n}|H_0)) \quad (24)$$

On this logarithmic scale, positive values indicate support for H_1 , negative values indicate support for H_0 , and 0 indicates no support for either hypothesis. By convention (Kass & Raftery, 1995), if $-0.5 < \log_{10}(BF_{10}) < 0.5$ then the result is considered inconclusive: there is not substantial evidence for either hypothesis. We used the BayesFactor package (Morey & Rouder, 2022) in R for these calculations.

One way to accomplish post-hoc comparisons would be to use a Bayesian analogue of the t-test to test for equality between each pair of profile means. However, this approach suffers from two problems. First, each Bayesian t-test would only use the data from the two latent profiles being compared. This omits data from the other profiles, which is useful for estimating the outcome variance. Also, when there are more than three or four profiles, reporting and interpreting the results of all pairwise tests would be cumbersome.

Instead, we used a different post-hoc analysis method that arrives at a similar result to pairwise comparisons without suffering the same drawbacks (see Rouder et al., 2012). Each pairwise comparison can have two conclusions: the means are either equal or unequal. What results from

the post-hoc analysis is a partitioning of the latent profiles into two or more sets, with equal profile means within each set and different means between sets. For a given number of latent profiles, we can enumerate all possible partitions (and thus all possible post-hoc analysis results) through a mathematical algorithm (Hutchinson, 1963). For example, with three latent profiles the post-hoc analysis could result in the following possible partitions:

$$\begin{aligned}\mu_1^{(y)} \neq \mu_2^{(y)}; \mu_1^{(y)} \neq \mu_3^{(y)}; \mu_2^{(y)} = \mu_3^{(y)} &\rightarrow \{1\}, \{2, 3\} \\ \mu_1^{(y)} = \mu_2^{(y)}; \mu_1^{(y)} \neq \mu_3^{(y)}; \mu_2^{(y)} \neq \mu_3^{(y)} &\rightarrow \{1, 2\}, \{3\} \\ \mu_1^{(y)} \neq \mu_2^{(y)}; \mu_1^{(y)} = \mu_3^{(y)}; \mu_2^{(y)} \neq \mu_3^{(y)} &\rightarrow \{1, 3\}, \{2\} \\ \mu_1^{(y)} \neq \mu_2^{(y)}; \mu_1^{(y)} \neq \mu_3^{(y)}; \mu_2^{(y)} \neq \mu_3^{(y)} &\rightarrow \{1\}, \{2\}, \{3\}\end{aligned}$$

where the notation {1, 2}, {3} indicates that profiles 1 and 2 form a set with one shared mean, while profile 3 has a different mean and thus forms another set by itself.

Thus, instead of conducting pairwise Bayesian t-tests, we can enumerate all possible partitions of the latent profiles and decide which one describes the data best by comparing their model evidences (the calculation is the same as the Bayesian ANOVA).

Effect Sizes

We also computed a simple effect size measure for the relationship between the neurocognitive latent profiles and outcome variables. This was an r^2 statistic similar to the one used with linear regression. First, we computed each individual's predicted value ($\hat{y}_i$) based on the fitted DPM-LPA model:

$$\hat{y}_i = \mu_{[z_i]}^{(y)} = \text{mean of participant } i\text{'s profile} \quad (25)$$

Then we computed r^2 as the proportional reduction in error:

$$r^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (26)$$

where $\bar{y}$ is the overall mean of y .

RESULTS

SIMULATIONS

DPM-LPA has similar performance to conventional LPA and superior performance to finite Bayesian LPA (see Supplemental Figure 1). DPM-LPA detected the correct number of latent profiles more often than conventional LPA (using either the BIC or BLRT) at a small sample size of $n = 250$ (DPM-LPA: 92% correct, conventional LPA with BIC: 82% correct, conventional LPA with BLRT: 72% correct). DPM-LPA was equally accurate to conventional LPA using the BLRT at larger sample sizes ($n = 500, 1000$), although the BIC was slightly more accurate (BIC: 98% correct, DPM-LPA/BLRT: 92–94% correct). Finite Bayesian LPA was less accurate than either DPM-LPA or conventional LPA, with the difference increasing at larger sample sizes.

SELECTING THE NUMBER OF PROFILES IN CONVENTIONAL LPA

Commonly used criteria (entropy reduction statistic, AIC, BIC, bootstrap likelihood ratio test) led to contradictory conclusions about the correct number of latent profiles in our sample. The entropy reduction statistic became worse (i.e., decreased) as the number of profiles increased (Figure 4a), falling to low levels (around 0.7) for models with more than 4 profiles. This suggested that the correct number of profiles is small. However, AIC, BIC, and the bootstrap likelihood test led to the opposite conclusion. AIC and BIC improved as the number of profiles increased (Figure 4b, note that AIC and BIC are reverse coded, i.e. multiplied by -1 , so larger values are better), showing that adding more profiles improved model fit. Similarly, the bootstrap likelihood ratio test reported a statistically significant improvement in model fit ($p < 0.01$) with each profile added up to 16 profiles.

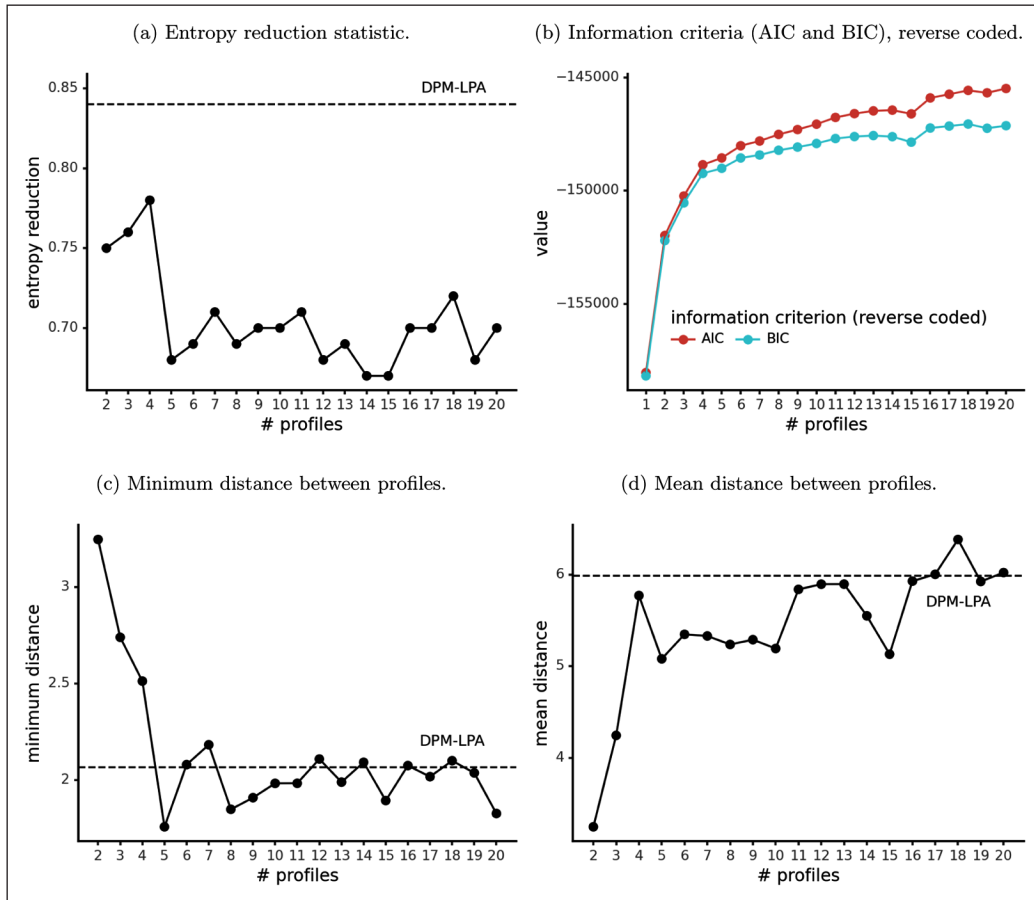


Figure 4 Comparisons of conventional LPA models with differing numbers of profiles (all variables are coded so that larger means better). Where applicable, a dashed horizontal line indicates the corresponding value from DPM-LPA.

COMPARING CONVENTIONAL LPA AND DPM-LPA: ENTROPY REDUCTION STATISTIC

DPM-LPA had a better entropy reduction statistic than any conventional LPA model (Figure 4a). In other words, DPM-LPA was more confident than conventional LPA in its classification of participants into latent profiles. This may be due to the relative distinctiveness of DPM-LPA's profiles, as described above.

COMPARING CONVENTIONAL LPA AND DPM-LPA: PROFILE DISTINCTIVENESS (MAHALANOBIS DISTANCE)

In conventional LPA, smaller models (2, 3, or 4 profiles) had a higher minimum pairwise distance between profiles than the larger ones (Figure 4c). In contrast, mean pairwise distance between profiles was lower for the very small conventional LPA models (2 or 3 profiles) than the larger ones (Figure 4d). In general, DPM-LPA was comparable or better than conventional LPA models with 5 or more profiles on both measures of distinctiveness (minimum and mean pairwise distance). Small conventional LPA models (2, 3, or 4 profiles) had better minimum distance than DPM-LPA, but worse mean distance. DPM-LPA (which discovered 9 latent profiles) was superior to the 9-profile conventional LPA model on both measures.

DESCRIPTION OF CONVENTIONAL LPA LATENT PROFILES

We selected the 4-profile model because its entropy reduction statistic was acceptable (0.78); larger models had unacceptably low entropy reduction statistics (0.67–0.72), despite their superior BIC/AIC and support from the BLRT. Based on these criteria, a four profile solution was selected (see Supplemental Figure 2): Profile 1 (56% of the sample) represented average neurocognition, Profile 2 (21% of the sample) represented above average neurocognition, Profile 3 (21% of the sample) represented below average neurocognition, and Profile 4 (2% of the sample) represented below average neurocognition with exceptionally low working memory.

DPM-LPA identified nine latent profiles. Figure 5 shows plots of the estimated means for each profile. Profile 1 represented average performance across neurocognitive components and encompassed 58% of the sample. Profile 2 (12% of the sample) represented above average performance across most neurocognitive components. Profile 3 (10% of the sample) represented below average performance across most neurocognitive components, with particularly low vocabulary and reading decoding but average response inhibition. Participants in Profile 4 (9% of the sample) had broadly above average neurocognitive function, with particularly good vocabulary and reading decoding. Participants in Profile 5 (5% of the sample) had below average working memory, recognition memory, and response inhibition. Profile 6 (3% of the sample) was characterized by low cognitive/attentional control. Profile 7 (2% of the sample) was characterized by above average processing speed, cognitive/attentional control, and spatial processing. Profiles 8 and 9 each contained 1% of the sample. Profile 8 showed generally poor neurocognitive functioning, while participants in Profile 9 had average performance on most tasks but exceptionally low working memory. Overall, profiles 2, 4, and 7 represented different patterns of above average neurocognitive function, while profiles 3, 5, 6, 8, and 9 represented different patterns of below average neurocognitive function. Importantly, each profile represented a unique distribution of neurocognitive scores.

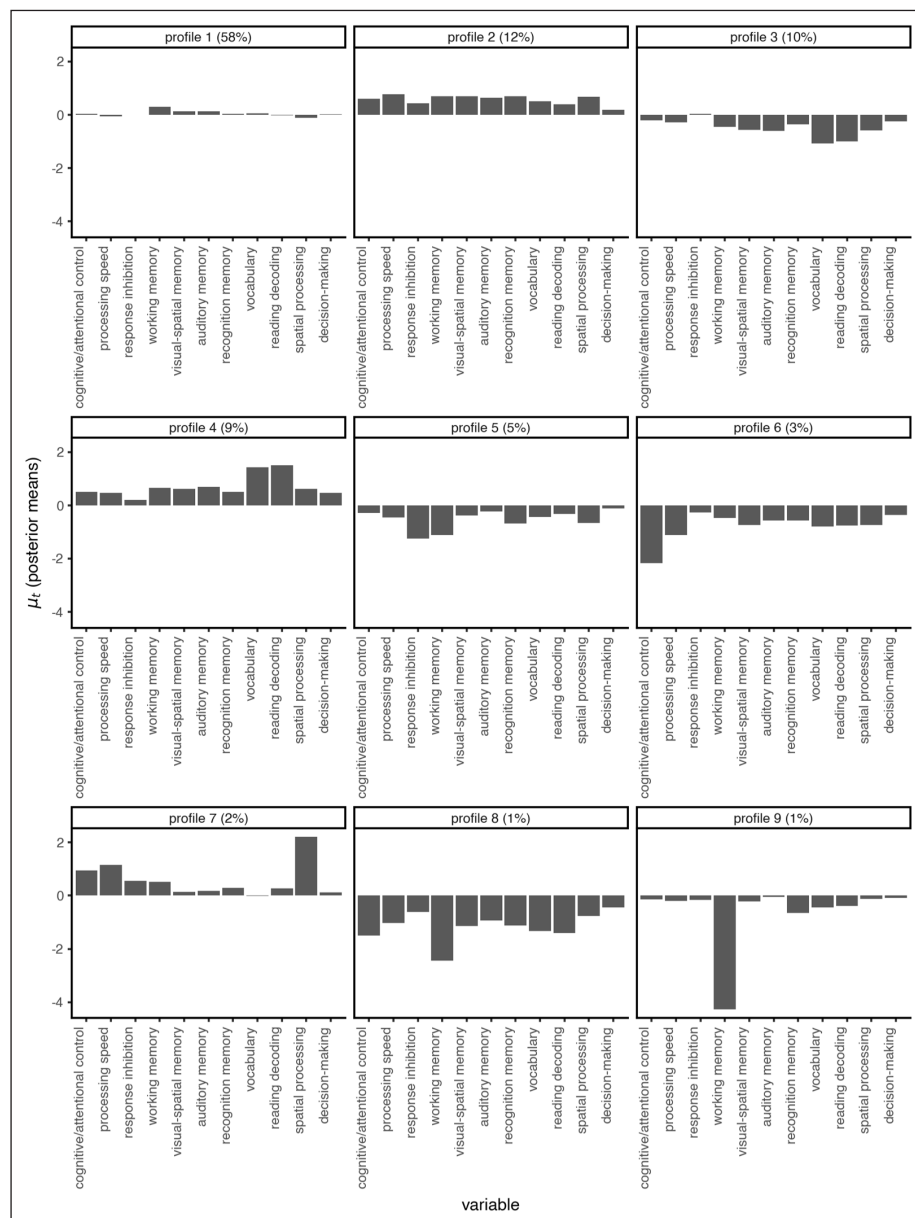


Figure 5 Estimated latent profile means (μ_t) from DPM-LPA.

See Figure 6 for plots of estimated profile means for all outcome variables and Table 2 for a summary of results. For comparison with DPM-LPA, we conducted a similar analysis using the 4-profile conventional LPA model (Supplemental Table 2). Results using conventional LPA were broadly similar to the DPM-LPA results reported below, but the DPM-LPA analysis provided a more nuanced description of the relationship between neurocognition and externalizing behaviors.

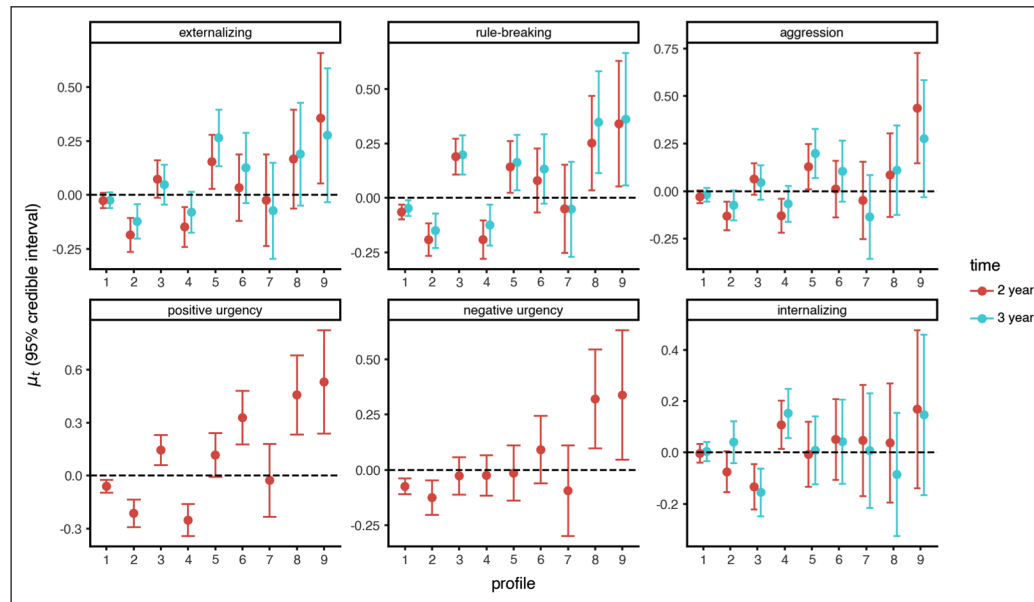


Figure 6 Estimated DPM-LPA latent profile outcome means (95% posterior credible interval).

(A) TWO YEAR FOLLOW-UP			
VARIABLE	$\log_{10}(BF_{10})$	r^2	POST-HOC RESULTS
externalizing	3.51	0.0090	$\mu_2^{(y)} = \mu_4^{(y)} < \mu_1^{(y)} = \mu_6^{(y)} = \mu_7^{(y)} < \mu_3^{(y)} = \mu_5^{(y)} = \mu_8^{(y)} = \mu_9^{(y)}$
rule-breaking	11.59	0.0167	$\mu_2^{(y)} = \mu_4^{(y)} < \mu_1^{(y)} = \mu_7^{(y)} < \mu_3^{(y)} = \mu_5^{(y)} = \mu_6^{(y)} = \mu_8^{(y)} = \mu_9^{(y)}$
aggression	1.22	0.0070	$\mu_2^{(y)} = \mu_4^{(y)} < \mu_1^{(y)} = \mu_6^{(y)} = \mu_7^{(y)} < \mu_3^{(y)} = \mu_5^{(y)} = \mu_8^{(y)} < \mu_9^{(y)}$
positive urgency	18.87	0.0230	$\mu_2^{(y)} = \mu_4^{(y)} < \mu_1^{(y)} = \mu_7^{(y)} < \mu_3^{(y)} = \mu_5^{(y)} < \mu_6^{(y)} = \mu_8^{(y)} = \mu_9^{(y)}$
negative urgency	-0.47	0.0052	NA (inconclusive)
internalizing	-1.57	NA	NA (no differences)
(B) THREE YEAR FOLLOW-UP			
VARIABLE	$\log_{10}(BF_{10})$	r^2	POST-HOC RESULTS
externalizing	2.11	0.0079	$\mu_2^{(y)} = \mu_4^{(y)} = \mu_7^{(y)} < \mu_1^{(y)} = \mu_3^{(y)} < \mu_5^{(y)} = \mu_6^{(y)} = \mu_8^{(y)} = \mu_9^{(y)}$
rule-breaking	8.01	0.0140	$\mu_2^{(y)} = \mu_4^{(y)} < \mu_1^{(y)} = \mu_7^{(y)} < \mu_3^{(y)} = \mu_5^{(y)} = \mu_6^{(y)} = \mu_8^{(y)} = \mu_9^{(y)}$
aggression	-0.85	NA	NA (no differences)
internalizing	-0.62	NA	NA (no differences)

Table 2 Summary of outcome analysis results. $\log_{10}(BF_{10})$ is the Bayes factor for the Bayesian ANOVA on a base-10 logarithmic scale. r^2 is the effect size. $\mu_j^{(y)}$ is the mean value of the relevant outcome variable for profile j .

Externalizing

Externalizing, Rule-Breaking, and Aggression

Overall externalizing behaviors at the two-year follow-up differed across latent profiles ($\log_{10}(BF_{10}) = 3.51$, $BF_{10} = 3.22 \times 10^3$, $r^2 = 0.0090$). Post-hoc tests showed that profiles 6 and 7 had equal means to profile 1 (average neurocognition), profiles 2 and 4 were related to lower levels of externalizing, while profiles 3, 5, 8, and 9 were related to higher levels of externalizing.

The rule-breaking and aggression subscales also differed across latent profiles (rule-breaking: $\log_{10}(BF_{10}) = 11.59$, $BF_{10} = 3.87 \times 10^{11}$, $r^2 = 0.0167$; aggression: $\log_{10}(BF_{10}) = 1.22$, $BF_{10} = 16.70$, $r^2 = 0.0070$). However, the relationship with neurocognition was much stronger for rule-breaking ($r^2 = 0.0167$) than aggression ($r^2 = 0.0070$). Post-hoc tests for rule-breaking revealed that profile 7 (posterior mean estimate: $\hat{\mu}_7^{(y)} = -0.05$) had the same mean as profile 1 (average neurocognition, $\hat{\mu}_1^{(y)} = -0.06$), profiles 2 ($\hat{\mu}_2^{(y)} = -0.19$) and 4 ($\hat{\mu}_4^{(y)} = -0.19$) had a lower mean level of rule-breaking, and profiles 3 ($\hat{\mu}_3^{(y)} = 0.19$), 5 ($\hat{\mu}_5^{(y)} = 0.14$), 6 ($\hat{\mu}_6^{(y)} = 0.08$), 8 ($\hat{\mu}_8^{(y)} = 0.25$), and 9 ($\hat{\mu}_9^{(y)} = 0.34$) showed higher mean levels of rule-breaking. For aggression, profiles 6 ($\hat{\mu}_6^{(y)} = 0.01$) and 7 ($\hat{\mu}_7^{(y)} = -0.05$) had the same mean as profile 1 ($\hat{\mu}_1^{(y)} = -0.03$), profiles 2 ($\hat{\mu}_2^{(y)} = 0.13$) and 4 ($\hat{\mu}_4^{(y)} = -0.13$) showed lower mean levels of aggression, profiles 3 ($\hat{\mu}_3^{(y)} = 0.06$), 5 ($\hat{\mu}_5^{(y)} = 0.13$), and 8 ($\hat{\mu}_8^{(y)} = 0.08$) showed higher mean aggression, and profile 9 ($\hat{\mu}_9^{(y)} = 0.44$) had the highest mean level of aggression.

At the three-year follow-up, the latent profiles were still related to overall externalizing behaviors ($\log_{10}(BF_{10}) = 2.11$, $BF_{10} = 1.30 \times 10^2$, $r^2 = 0.0079$) and rule-breaking ($\log_{10}(BF_{10}) = 8.01$, $BF_{10} = 1.03 \times 10^8$, $r^2 = 0.0140$), but not aggression ($\log_{10}(BF_{10}) = -0.85$, $BF_{10} = 0.14$). Post-hoc results for rule-breaking were the same as those for the two-year follow-up described above. However, the post-hoc results for overall externalizing changed at the three-year follow-up. Profiles 1 ($\hat{\mu}_1^{(y)} = -0.02$) and 3 ($\hat{\mu}_3^{(y)} = 0.05$) shared the same mean externalizing, profiles 2 ($\hat{\mu}_2^{(y)} = -0.12$), 4 ($\hat{\mu}_4^{(y)} = -0.08$), and 7 ($\hat{\mu}_7^{(y)} = -0.07$) had a lower mean externalizing, and profiles 5 ($\hat{\mu}_5^{(y)} = 0.26$), 6 ($\hat{\mu}_6^{(y)} = 0.13$), 8 ($\hat{\mu}_8^{(y)} = 0.19$), and 9 ($\hat{\mu}_9^{(y)} = 0.28$) had a higher mean externalizing.

Positive and Negative Urgency

Positive urgency differed across neurocognitive latent profiles ($\log_{10}(BF_{10}) = 18.87$, $BF_{10} = 7.46 \times 10^{18}$, $r^2 = 0.0230$). Post-hoc tests showed that profile 7 ($\hat{\mu}_7^{(y)} = -0.03$) had the same positive urgency mean as profile 1 ($\hat{\mu}_1^{(y)} = -0.06$), profile 2 ($\hat{\mu}_2^{(y)} = -0.21$) and profile 4 ($\hat{\mu}_4^{(y)} = -0.25$) had a lower positive urgency mean, profile 3 ($\hat{\mu}_3^{(y)} = 0.14$) and profile 5 ($\hat{\mu}_5^{(y)} = 0.12$) had a higher positive urgency mean, while profiles 6 ($\hat{\mu}_6^{(y)} = 0.33$), 8 ($\hat{\mu}_8^{(y)} = 0.46$), and 9 ($\hat{\mu}_9^{(y)} = 0.53$) had the highest positive urgency mean. The Bayes factor testing the relationship between neurocognitive profiles and negative urgency was indecisive ($\log_{10}(BF_{10}) = -0.47$, $BF_{10} = 0.34$, $r^2 = 0.0052$); if there is a relationship, it is weaker than for positive urgency (negative urgency: $r^2 = 0.0052$ vs. positive urgency: $r^2 = 0.0230$).

Interim Summary

Altogether, profiles 2 and 4 seem to capture a group of adolescents typified by above average neurocognitive performance and good behavioral regulation capacities across the two-and-three-year data (consistently had the lowest levels of overall externalizing behaviors, rule-breaking, and aggression). In contrast, profiles 3, 5, 6, 8 and 9 showed below average performance across most neurocognitive measures, and overall had higher than average levels of externalizing behaviors. Profile 3 included adolescents who struggled the most with tasks that rely on vocabulary and reading skills and display higher than average levels of externalizing, rule-breaking, and aggression (although their level of externalizing declined to the same level as profile 1, the average group, at the three-year follow-up). Profile 5 captured adolescents who showed relatively poorer executive functions (working memory and response inhibition), which related to their externalizing behaviors. Adolescents in profile 6 showed a particular pattern of neurocognitive scores that indicates that they found it challenging to engage cognitive/attentional control, which was related to higher scores on rule-breaking and positive urgency but not aggression. Finally, profiles 8 and 9 appear to capture the most severe cases of general neurocognitive problems, particularly in terms of working memory, and these groups showed higher levels externalizing behaviors across measures and time.

Internalizing

Sensitivity analysis showed that the neurocognitive profiles were not related to the CBCL internalizing scale at either the two-year-follow-up ($\log_{10}(BF_{10}) = -1.57$, $BF_{10} = 0.03$) or the three-year-follow-up ($\log_{10}(BF_{10}) = -0.62$, $BF_{10} = 0.24$). This suggest that the differences in neurocognitive patterns within-persons are specific to externalizing-related outcomes (see also Chaku et al., 2022).

The overarching goal of this work was to develop and test a novel approach to study the relationship between individual differences in neurocognitive functioning and externalizing behaviors in adolescents. Our findings highlight the advantages and power of a novel non-parametric Bayesian approach to LPA, called DPM-LPA, for untangling subgroups of adolescents based on an extensive and diverse set of neurocognitive metrics, even when groups overlap in the patterns of observable externalizing behaviors.

COMPARING LPA METHODS

Consistent with previous work (e.g. Chaku et al., 2022; Wendt et al., 2019), different selection criteria gave different answers about the correct number of profiles in conventional LPA (Figure 4). Model fit (measured by AIC and the bootstrap likelihood ratio test) improved by increasing the number of profiles, but this came at the cost of having highly similar, redundant profiles and hence lower classification certainty (measured by the entropy reduction statistic). This illustrates the tradeoff in conventional LPA between model fit and model interpretability.

In contrast, DPM-LPA automatically decided how many latent profiles to include. Compared to conventional LPA models with equal flexibility, DPM-LPA produced less similar profiles and classified participants into profiles with greater certainty (i.e., had better entropy reduction, Figure 4a). This can be explained by DPM-LPA's non-parametric Bayesian inference process. Fitting conventional LPA only involves maximizing the likelihood, which rewards the fitting algorithm for using all available latent profiles to classify people even if some of them end up being very similar (Figure 4d). However, DPM-LPA takes both the likelihood and prior distribution into account. The prior favors a small number of profiles that actually contain participants, leaving unneeded profiles empty. Thus, DPM-LPA has the flexibility to infer a large number of latent profiles if it needs to, but does not tend to infer redundant, highly similar profiles, in the way that conventional LPA does. Simulations provide further support for DPM-LPA (see Supplemental Material): it compares well with conventional LPA in terms of detecting the correct number of latent profiles, and may be the superior method at small sample sizes ($n = 250$).

NEUROCOGNITIVE LATENT PROFILES AND EXTERNALIZING BEHAVIORS

In addition to methodological superiority over conventional LPA, DPM-LPA also provides a general approach to generate unique insights into how the same externalizing behaviors can be associated with differences in patterns of neurocognition across subgroups. Broadly speaking, in the present study, the profiles showing more externalizing behaviors were typified by below average neurocognitive functioning, and those characterized by above average neurocognitive performance showed the opposite pattern. However, most importantly, we were able to further pinpoint subgroup-specific patterns reflecting issues with subcomponents of executive functions (working memory, inhibition) and language processing in the latent profiles showing relatively higher levels of externalizing.

The presence of subgroups predominantly showing reduced working memory capacity is in line with the proposal that executive function plays a key role in the development of reduced behavioral regulation during childhood (Moffitt, 1993; Zelazo, 2020). Notably, we also identified subgroups with equally worse performance on two subcomponents of executive functions, working memory and inhibition (profile 5), and two out of three subgroups showing poorer working memory also presented with worse inhibition (profiles 8 and 9). This pattern suggests a possible interaction between these two subcomponents of executive functions, which converges with prior evidence in adults indicating that cognitive inhibition accounts for part of the variability in age-related working memory (Borella, Carretti, & De Beni, 2008). In sum, these findings further highlight the sensitivity of the DPM-LPA for detecting latent profiles based on fine-grained differences in neurocognitive functions, perhaps even providing sufficient sensitivity to disentangle subgroups in which the interaction of different subcomponents of neurocognitive functions could turn out to be a key etiological factor. Future studies could explore this possibility further.

Our findings also highlight the existence of a group of adolescents with externalizing proneess and below average general neurocognitive functioning that especially struggle with language development (profile 3). Prior work has shown that language skills play a particularly important role in the development of externalizing behaviors, with one study indicating that having poorer language abilities is a temporally stable within-person predictor of having more externalizing problems later in childhood (Petersen & LeBeau, 2021). The importance of identifying which adolescents present with poor language abilities becomes particularly salient in the context of intervening on externalizing behaviors, as many interventions are delivered verbally and thus place a burden on individuals' language skills. It may be necessary to support these individuals through alternative strategies that take into consideration their neurocognitive difficulties.

Importantly, our results suggest that the association between neurocognition and externalizing behaviors varies depending on the subgroup of neurocognition and subtype of externalizing behaviors. For example, we saw divergence in the association between variability in neurocognition and engagement in impulsive behavior when feeling positive versus negative emotions. Below average neurocognitive profiles all related to higher positive urgency (albeit to different degrees). There was little evidence of differences in neurocognition related to negative urgency. It is possible that for this age-group, emotion-relevant impulsivity, particularly positive emotions, relates to neurocognitive vulnerability. It also is possible that the association between neurocognitive difficulties and negative urgency emerges at more extreme levels of impulsive behavior (Stautz & Cooper, 2013). Far more work is needed to examine if neurocognitive associations differentially relate to positive and negative urgency (Johnson, Elliott, & Carver, 2020; Johnson, Tharp, Peckham, Sanchez, & Carver, 2016; Littlefield, Stevens, Ellingson, King, & Jackson, 2016). As another example, we saw a lack of temporal stability in the association between neurocognitive profiles and aggressive behavior measured at timepoints two (age 11–12) and three (age 12–13). This finding is consistent with developmental research which points out that aggression and rule-breaking represent different dimensions within the context of broad externalizing behaviors, each with a separate pattern of (neuro)biological correlates and developmental timeline (Burt, 2013; Harden et al., 2015). While aggression decreases as children age, the tendency to break rules increases over time and into adolescence (Bongers, Koot, Van Der Ende, & Verhulst, 2004). Follow-up studies in future waves of the ABCD data collection could try to elucidate the developmental changes in neurocognition and behavioral outcomes.

LIMITATIONS

This study had several limitations which can be addressed in future work. First, based on the ABCD two-year follow-up neurocognitive data we find that DPM-LPA is superior to conventional LPA. Examining the robustness of this claim with other datasets, varying in sample type, measures of neurocognition, and the number of variables will be an important next step.

Second, while the neurocognitive battery used at the two-year-follow-up is quite extensive, it did not tap all neurocognitive functions (e.g., cognitive flexibility) or include tasks with more ecologically-valid stimuli (Conley & Baskin-Sommers, 2023; Young, Frankenhuys, DelPriore, & Ellis, 2022). Therefore, the pattern of neurocognitive functions may shift when considering additional cognitive and task-related factors. Related, adolescence is a time of rapid neurocognitive and behavioral development. Therefore, we may obtain very different neurocognitive latent profiles at an older or younger age. Future work should investigate the stability or change in neurocognitive latent profiles and their relationship to externalizing behaviors.

Third, our DPM-LPA software is limited in several ways and can be improved upon in later versions. The software can currently only model indicator variables using normal distributions. However, many psychological applications of LPA involve binary or categorical indicators, or indeed a combination of variable types. Implementing greater flexibility in the allowable distributions of indicators would improve the applicability of DPM-LPA (e.g. Qiu et al., 2023).

Finally, our simulations comparing DPM-LPA to conventional LPA and finite Bayesian LPA were limited by the empirical focus of the current study. Simulations for evaluating statistical methods

require substantial computational resources, and are often full studies in and of themselves (e.g. Tein et al., 2013). We only simulated one possible scenario (10 indicator variables, 5 latent profiles, equal number of participants in each latent profile, profile means separated by 1.5 standard deviations) out of many possible ones. We also only generated 50 simulated data sets in each condition: more replications could improve our estimates of each model/method's accuracy. Further, detecting the correct number of latent profiles is not the only criterion by which to compare methods: accuracy of assigning participants to the correct latent profiles and accuracy of estimating profile means also are important. Conducting extensive simulations that address these limitations would be a useful direction for future research.

CONCLUSIONS

This paper presents a novel implementation of DPM-LPA, a non-parametric Bayesian approach to conducting LPA, which can detect latent profiles from relatively large amounts of indicator variables and offers a possible solution for dealing with the tradeoff between interpretability and model fit that plagues conventional LPA. We showed that DPM-LPA can be used to better understand how externalizing behaviors seen across individuals relate to differences in neurocognitive subcomponents by detecting latent subgroups of individuals with similar neurocognition. Our study marks a step towards addressing the challenge of finding novel ways to use data on neurocognitive functioning to better describe the individual (Brazil et al., 2018; Ruiz et al., 2023). Such advances are particularly relevant given that current call for person-centered diagnostics and tailored interventions for individuals showing these costly behaviors.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Supplemental material.** Comparison of participants with complete and incomplete data, simulations, conventional LPA outcome analysis. DOI: <https://doi.org/10.5334/cpsy.112.s1>

FUNDING INFORMATION

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive DevelopmentSM (ABCD) Study (<https://abcdstudy.org>), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9–10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at <https://abcdstudy.org/federal-partners.html>. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. Funding for this work also came from R21DA057592-01 and a WTI Innovation Grant (Yale).

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Conceptualization: S.P., I.Y., S.G.R., I.A.B., A.B.-S.; Methodology: S.P., I.A.B., A.B.-S.; Formal Analysis: S.P., A.B.-S.; Visualization: S.P.; Writing—Original Draft Preparation: S.P., I.A.B., A.B.-S.; Writing—Review & Editing: S.P., I.Y., S.G.R., I.A.B., A.B.-S.; Supervision: I.A.B., A.B.-S.

AUTHOR AFFILIATIONS

Sam Paskewitz  orcid.org/0000-0003-4886-6199

Department of Psychology, Yale University, New Haven CT, US

Inti A. Brazil  orcid.org/0000-0001-5824-0902

Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands;
Forensic Psychiatric Centre Pompestichting, Nijmegen, The Netherlands

Ilker Yildirim  orcid.org/0000-0002-9072-0938

Department of Psychology, Yale University, New Haven CT, US

Sonia Ruiz  orcid.org/0000-0002-3681-946X

Department of Psychology, Yale University, New Haven CT, US

Arielle Baskin-Sommers  orcid.org/0000-0001-6773-0508

Department of Psychology, Yale University, New Haven CT, US

REFERENCES

- Achenbach, T. M.** (2009). *The Achenbach system of empirically based assessment (ASEBA): Development, findings, theory, and applications*. University of Vermont, Research Center for Children, Youth, & Families.
- Acker, W., & Acker, C.** (1982). *Bexley Maudsley automated processing screening and Bexley Maudsley category sorting test manual*. NFER-Nelson Publishing.
- Akaike, H.** (1974, December). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. Retrieved 2023-10-08, from <http://ieeexplore.ieee.org/document/1100705>. DOI: <https://doi.org/10.1109/TAC.1974.1100705>
- Barch, D. M., Albaugh, M. D., Avenevoli, S., Chang, L., Clark, D. B., Glantz, M. D., ... Sher, K. J.** (2018, August). Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. *Developmental Cognitive Neuroscience*, 32, 55–66. Retrieved 2023-03-09, from <https://linkinghub.elsevier.com/retrieve/pii/S1878929317300683>. DOI: <https://doi.org/10.1016/j.dcn.2017.10.010>
- Beauchaine, T. P., & McNulty, T.** (2013). Comorbidities and continuities as ontogenic processes: Toward a developmental spectrum model of externalizing psychopathology. *Development and Psychopathology*, 25(4pt2), 1505–1528. (Publisher: Cambridge University Press). DOI: <https://doi.org/10.1017/S0954579413000746>
- Blackwell, D., & MacQueen, J. B.** (1973). Ferguson distributions via Pólya urn schemes. *The Annals of Statistics*, 1(2), 353–355. (Publisher: Institute of Mathematical Statistics). DOI: <https://doi.org/10.1214/aos/1176342372>
- Blei, D. M., & Jordan, M. I.** (2006, March). Variational inference for Dirichlet process mixtures. *Bayesian Analysis*, 1(1), 121–143. Retrieved 2020-11-12, from <http://projecteuclid.org/euclid.ba/1340371077>. DOI: <https://doi.org/10.1214/06-BA104>
- Blei, D. M., Kucukelbir, A., & McAuliffe, J. D.** (2017, April). Variational Inference: A Review for Statisticians. *Journal of the American Statistical Association*, 112(518), 859–877. Retrieved 2022-08-21, from <https://www.tandfonline.com/doi/full/10.1080/01621459.2017.1285773>. DOI: <https://doi.org/10.1080/01621459.2017.1285773>
- Bongers, I. L., Koot, H. M., Van Der Ende, J., & Verhulst, F. C.** (2004). Developmental trajectories of externalizing behaviors in childhood and adolescence. *Child Development*, 75(5), 1523–1537. (Publisher: Wiley Online Library). DOI: <https://doi.org/10.1111/j.1467-8624.2004.00755.x>
- Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C.** (2008, July). Predicting young adult social functioning from developmental trajectories of externalizing behaviour. *Psychological medicine*, 38(7), 989–999. (Place: England). DOI: <https://doi.org/10.1017/S0033291707002309>
- Borella, E., Carretti, B., & De Beni, R.** (2008). Working memory and inhibition across the adult life-span. *Acta Psychologica*, 128(1), 33–44. (Publisher: Elsevier). DOI: <https://doi.org/10.1016/j.actpsy.2007.09.008>
- Brand, M., Fujiwara, E., Borsutzky, S., Kalbe, E., Kessler, J., & Markowitsch, H. J.** (2005, May). Decision-Making Deficits of Korsakoff Patients in a New Gambling Task With Explicit Rules: Associations With Executive Functions. *Neuropsychology*, 19(3), 267–277. Retrieved 2023-04-14, from <http://doi.apa.org/getdoi.cfm?doi=10.1037/0894-4105.19.3.267>. DOI: <https://doi.org/10.1037/0894-4105.19.3.267>
- Brazil, I. A., van Dongen, J. D., Maes, J. H., Mars, R., & Baskin-Sommers, A. R.** (2018). Classification and treatment of antisocial individuals: From behavior to biocognition. *Neuroscience & Biobehavioral Reviews*, 91, 259–277. (Publisher: Elsevier). DOI: <https://doi.org/10.1016/j.neubiorev.2016.10.010>
- Burt, S.** (2013). Do etiological influences on aggression overlap with those on rule breaking? A meta-analysis. *Psychological Medicine*, 43(9), 1801–1812. (Publisher: Cambridge University Press). DOI: <https://doi.org/10.1017/S0033291712001894>

- Casey, B., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., ... Dale, A. M. (2018, August). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, 32, 43–54. Retrieved 2023-03-09, from <https://linkinghub.elsevier.com/retrieve/pii/S1878929317301214>. DOI: <https://doi.org/10.1016/j.dcn.2018.03.001>
- Chaku, N., Barry, K., Fowle, J., & Hoyt, L. T. (2022, November). Understanding patterns of heterogeneity in executive functioning during adolescence: Evidence from population-level data. *Developmental Science*, 25(6), e13256. Retrieved 2023-10-10, from <https://onlinelibrary.wiley.com/doi/10.1111/desc.13256>. DOI: <https://doi.org/10.1111/desc.13256>
- Clark, D. B., Fisher, C. B., Bookheimer, S., Brown, S. A., Evans, J. H., Hopfer, C., ... others. (2018). Biomedical ethics and clinical oversight in multisite observational neuroimaging studies with children and adolescents: The ABCD experience. *Developmental Cognitive Neuroscience*, 32, 143–154. (Publisher: Elsevier). DOI: <https://doi.org/10.1016/j.dcn.2017.06.005>
- Conley, M., & Baskin-Sommers, A. (2023). Development in uncertain contexts: An ecologically informed approach to understanding decision-making during adolescence. *Cognitive, Affective, & Behavioral Neuroscience*, 23(3), 739–745. (Publisher: Springer). DOI: <https://doi.org/10.3758/s13415-023-01067-7>
- Cyders, M. A., Smith, G. T., Spillane, N. S., Fischer, S., Annus, A. M., & Peterson, C. (2007). Integration of impulsivity and positive mood to predict risky behavior: development and validation of a measure of positive urgency. *Psychological Assessment*, 19(1), 107. (Publisher: American Psychological Association). DOI: <https://doi.org/10.1037/1040-3590.19.1.107>
- Danielson, M. L., Bitsko, R. H., Holbrook, J. R., Charania, S. N., Claussen, A. H., McKeown, R. E., ... Flory, K. (2021, June). Community-Based Prevalence of Externalizing and Internalizing Disorders among School-Aged Children and Adolescents in Four Geographically Dispersed School Districts in the United States. *Child Psychiatry and Human Development*, 52(3), 500–514. (Place: United States). DOI: <https://doi.org/10.1007/s10578-020-01027-z>
- De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, 131(4), 483. (Publisher: American Psychological Association). DOI: <https://doi.org/10.1037/0033-2909.131.4.483>
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977, September). Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1), 1–22. Retrieved 2020-01-21, from <http://doi.wiley.com/10.1111/j.2517-6161.1977.tb01600.x>. DOI: <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>
- Fairchild, G., Hawes, D. J., Frick, P. J., Copeland, W. E., Odgers, C. L., Franke, B., ... De Brito, S. A. (2019). Conduct disorder. *Nature Reviews Disease Primers*, 5(1), 43. (Publisher: Nature Publishing Group UK London). DOI: <https://doi.org/10.1038/s41572-019-0095-y>
- Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *The annals of statistics*, 209–230. (Publisher: JSTOR). DOI: <https://doi.org/10.1214/aos/1176342360>
- Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R., Heeringa, S., ... Zahs, D. (2018). Recruiting the ABCD sample: Design considerations and procedures. *Developmental Cognitive Neuroscience*, 32, 16–22. (Publisher: Elsevier). DOI: <https://doi.org/10.1016/j.dcn.2018.04.004>
- Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. Wiley New York.
- Harden, K. P., Patterson, M. W., Briley, D. A., Engelhardt, L. E., Kretsch, N., Mann, F. D., ... Tucker-Drob, E. M. (2015). Developmental changes in genetic and environmental influences on rule-breaking and aggression: age and pubertal development. *Journal of Child Psychology and Psychiatry*, 56(12), 1370–1379. (Publisher: Wiley Online Library). DOI: <https://doi.org/10.1111/jcpp.12419>
- Honrath, P., Kohls, G., Moore, T. M., Dogan, I., Gur, R. C., Habel, U., ... Gur, R. E. (2021). Investigating neurocognitive functioning in youths with externalizing disorders from the Philadelphia Neurodevelopmental Cohort. *Journal of Adolescent Health*, 69(1), 100–107. (Publisher: Elsevier). DOI: <https://doi.org/10.1016/j.jadohealth.2020.10.029>
- Hutchinson, G. (1963, October). Partitioning algorithms for finite sets. *Communications of the ACM*, 6(10), 613–614. Retrieved 2023-12-12, from <https://dl.acm.org/doi/10.1145/367651.367661>. DOI: <https://doi.org/10.1145/367651.367661>
- Ishwaran, H., & James, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association*, 96(453), 161–173. (Publisher: Taylor & Francis). DOI: <https://doi.org/10.1198/016214501750332758>
- Jefferys, W. H., & Berger, J. O. (1992). Ockham's Razor and Bayesian Analysis. *American Scientist*, 80(1), 64–72.
- Jernigan, T. L., Brown, S. A., & Dowling, G. J. (2018). The Adolescent Brain Cognitive Development Study. *Journal of Research on Adolescence*, 28(1), 154. (Publisher: NIH Public Access). DOI: <https://doi.org/10.1111/jora.12374>

- Johnson, S. L., Elliott, M. V., & Carver, C. S.** (2020, February). Impulsive Responses to Positive and Negative Emotions: Parallel Neurocognitive Correlates and Their Implications. *Biological Psychiatry*, 87(4), 338–349. Retrieved 2023-10-18, from <https://linkinghub.elsevier.com/retrieve/pii/S0006322319316312>. DOI: <https://doi.org/10.1016/j.biopsych.2019.08.018>
- Johnson, S. L., Tharp, J. A., Peckham, A. D., Sanchez, A. H., & Carver, C. S.** (2016, August). Positive urgency is related to difficulty inhibiting prepotent responses. *Emotion*, 16(5), 750–759. Retrieved 2023-10-18, from <http://doi.apa.org/getdoi.cfm?doi=10.1037/emo0000182>. DOI: <https://doi.org/10.1037/emo0000182>
- Kass, R. E., & Raftery, A. E.** (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773–795. (Publisher: Taylor & Francis). DOI: <https://doi.org/10.1080/01621459.1995.10476572>
- Langeheine, R., Pannekoek, J., & Van de Pol, F.** (1996). Bootstrapping goodness-of-fit measures in categorical data analysis. *Sociological Methods & Research*, 24(4), 492–516. (Publisher: SAGE PERIODICALS PRESS). DOI: <https://doi.org/10.1177/0049124196024004004>
- Li, Y., Lord-Bessen, J., Shiyko, M., & Loeb, R.** (2018). Bayesian latent class analysis tutorial. *Multivariate Behavioral Research*, 53(3), 430–451. (Publisher: Taylor & Francis). DOI: <https://doi.org/10.1080/00273171.2018.1428892>
- Littlefield, A. K., Stevens, A. K., Ellingson, J. M., King, K. M., & Jackson, K. M.** (2016). Changes in negative urgency, positive urgency, and sensation seeking across adolescence. *Personality and Individual Differences*, 90, 332–337. (Publisher: Elsevier). DOI: <https://doi.org/10.1016/j.paid.2015.11.024>
- Logan, G. D.** (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. (Publisher: Academic Press). DOI: <https://doi.org/10.1016/j.dcn.2018.02.006>
- Luciana, M., Bjork, J., Nagel, B., Barch, D., Gonzalez, R., Nixon, S., & Banich, M.** (2018, August). Adolescent neurocognitive development and impacts of substance use: Overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery. *Developmental Cognitive Neuroscience*, 32, 67–79. Retrieved 2023-03-09, from <https://linkinghub.elsevier.com/retrieve/pii/S1878929317302384>. DOI: <https://doi.org/10.1016/j.dcn.2018.02.006>
- Moffitt, T. E.** (1993, October). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological Review*, 100(4), 674–701. (Place: United States). DOI: <https://doi.org/10.1037//0033-295X.100.4.674>
- Morey, R. D., & Rouder, J. N.** (2022). *BayesFactor: Computation of Bayes Factors for Common Designs*. Retrieved from <https://CRAN.R-project.org/package=BayesFactor>.
- Ogilvie, J. M., Stewart, A. L., Chan, R. C. K., & Shum, D. H. K.** (2011, November). Neuropsychological Measures of Executive Function and Antisocial Behavior: a Meta-Analysis. *Criminology*, 49(4), 1063–1107. Retrieved 2023-09-27, from <https://onlinelibrary.wiley.com/doi/10.1111/j.1745-9125.2011.00252.x>. DOI: <https://doi.org/10.1111/j.1745-9125.2011.00252.x>
- Petersen, I. T., & LeBeau, B.** (2021). Language ability in the development of externalizing behavior problems in childhood. *Journal of Educational Psychology*, 113(1), 68. (Publisher: American Psychological Association). DOI: <https://doi.org/10.1037/edu0000461>
- Qiu, M., Paganin, S., Ohn, I., & Lin, L.** (2023). Bayesian nonparametric latent class analysis for different item types. *Multivariate Behavioral Research*, 58(1), 156–157. (Publisher: Taylor & Francis). DOI: <https://doi.org/10.1080/00273171.2022.2160958>
- Rosenberg, J. M., Beymer, P. N., Anderson, D. J., Van Lissa, C., & 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. DOI: <https://doi.org/10.21105/joss.00978>
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M.** (2012, October). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56(5), 356–374. Retrieved 2023-07-07, from <https://linkinghub.elsevier.com/retrieve/pii/S0022249612000806>. DOI: <https://doi.org/10.1016/j.jmp.2012.08.001>
- Ruiz, S. G., Brazil, I. A., & Baskin-Sommers, A.** (2023). Distinct neurocognitive fingerprints reflect differential associations with risky and impulsive behavior in a neurotypical sample. *Scientific Reports*, 13(1), 11782. (Publisher: Nature Publishing Group UK London). DOI: <https://doi.org/10.1038/s41598-023-38991-0>
- Schwarz, G.** (1978). Estimating the dimension of a model. *The Annals of Statistics*, 461–464. (Publisher: JSTOR). DOI: <https://doi.org/10.1214/aos/1176344136>
- Scrucca, L., Fraley, C., Murphy, T. B., & Raftery, A. E.** (2023). *Model-Based Clustering, Classification, and Density Estimation Using mclust in R*. Chapman and Hall/CRC. Retrieved from <https://mclust-org.github.io/book/>. DOI: <https://doi.org/10.1201/9781003277965>
- Shepherd, J., & Farrington, D.** (2003, April). The impact of antisocial lifestyle on health. *BMJ (Clinical research ed.)*, 326(7394), 834–835. (Place: England). DOI: <https://doi.org/10.1136/bmj.326.7394.834>
- Stautz, K., & Cooper, A.** (2013). Impulsivity-related personality traits and adolescent alcohol use: A meta-analytic review. *Clinical Psychology Review*, 33(4), 574–592. (Publisher: Elsevier). DOI: <https://doi.org/10.1016/j.cpr.2013.03.003>

- Strauss, E., Sherman, E. M., & Spreen, O.** (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York, NY: Oxford University Press.
- Taylor, B. K., Frenzel, M. R., Eastman, J. A., Wiesman, A. I., Wang, Y.-P., Calhoun, V. D., ... Wilson, T. W.** (2022). Reliability of the NIH toolbox cognitive battery in children and adolescents: a 3-year longitudinal examination. *Psychological Medicine*, 52(9), 1718–1727. (Publisher: Cambridge University Press). DOI: <https://doi.org/10.1017/S0033291720003487>
- Tein, J.-Y., Coxe, S., & Cham, H.** (2013, October). Statistical Power to Detect the Correct Number of Classes in Latent Profile Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 20(4), 640–657. Retrieved 2024-04-16, from <http://www.tandfonline.com/doi/abs/10.1080/10705511.2013.824781>. DOI: <https://doi.org/10.1080/10705511.2013.824781>
- Thompson, W. K., Barch, D. M., Bjork, J. M., Gonzalez, R., Nagel, B. J., Nixon, S. J., & Luciana, M.** (2019, April). The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: Findings from the ABCD study's baseline neurocognitive battery. *Developmental Cognitive Neuroscience*, 36, 100606. Retrieved 2023-09-27, from <https://linkinghub.elsevier.com/retrieve/pii/S1878929318301622>. DOI: <https://doi.org/10.1016/j.dcn.2018.12.004>
- Timmermans, M., van Lier, P. A. C., & Koot, H. M.** (2008, April). Which forms of child/adolescent externalizing behaviors account for late adolescent risky sexual behavior and substance use? *Journal of Child Psychology and Psychiatry*, 49(4), 386–394. (Place: England). DOI: <https://doi.org/10.1111/j.1469-7610.2007.01842.x>
- Van der Ende, J., Verhulst, F. C., & Tiemeier, H.** (2016, August). The bidirectional pathways between internalizing and externalizing problems and academic performance from 6 to 18 years. *Development and Psychopathology*, 28(3), 855–867. (Place: United States). DOI: <https://doi.org/10.1017/S0954579416000353>
- Van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., ... others.** (2021). Bayesian statistics and modelling. *Nature Reviews: Methods Primers*, 1(1), 1. (Publisher: Nature Publishing Group UK London). DOI: <https://doi.org/10.1038/s43586-020-00001-2>
- Van Goozen, S. H., Langley, K., & Hobson, C. W.** (2022). Childhood antisocial behavior: A neurodevelopmental problem. *Annual Review of Psychology*, 73, 353–377. (Publisher: Annual Reviews). DOI: <https://doi.org/10.1146/annurev-psych-052621-045243>
- Wang, M.-C., Deng, Q., Bi, X., Ye, H., & Yang, W.** (2017). Performance of the entropy as an index of classification accuracy in latent profile analysis: A Monte Carlo simulation study. *Acta Psychologica Sinica*. (Publisher: Science Press). DOI: <https://doi.org/10.3724/SP.J.1041.2017.01473>
- Wendt, L. P., Wright, A. G. C., Pilkonis, P. A., Nolte, T., Fonagy, P., Montague, P. R., ... Zimmermann, J.** (2019, November). The latent structure of interpersonal problems: Validity of dimensional, categorical, and hybrid models. *Journal of Abnormal Psychology*, 128(8), 823–839. Retrieved 2023-11-28, from <http://doi.apa.org/getdoi.cfm?doi=10.1037/abn0000460>. DOI: <https://doi.org/10.1037/abn0000460>
- White, A., & Murphy, T. B.** (2014). BayesLCA: An R package for Bayesian latent class analysis. *Journal of Statistical Software*, 61, 1–28. DOI: <https://doi.org/10.18637/jss.v061.i13>
- Whiteside, S. P., Lynam, D. R., Miller, J. D., & Reynolds, S. K.** (2005). Validation of the UPPS impulsive behaviour scale: a four-factor model of impulsivity. *European Journal of Personality*, 19(7), 559–574. London, England: SAGE Publications Sage UK. DOI: <https://doi.org/10.1002/per.556>
- Yang, Y., Shields, G. S., Zhang, Y., Wu, H., Chen, H., & Romer, A. L.** (2022, November). Child executive function and future externalizing and internalizing problems: A meta-analysis of prospective longitudinal studies. *Clinical Psychology Review*, 97, 102194. Retrieved 2023-09-27, from <https://linkinghub.elsevier.com/retrieve/pii/S0272735822000794>. DOI: <https://doi.org/10.1016/j.cpr.2022.102194>
- Young, E. S., Frankenhuys, W. E., DelPriore, D. J., & Ellis, B. J.** (2022). Hidden talents in context: Cognitive performance with abstract versus ecological stimuli among adversity-exposed youth. *Child Development*, 93(5), 1493–1510. (Publisher: Wiley Online Library). DOI: <https://doi.org/10.1111/cdev.13766>
- Zelazo, P. D.** (2020). Executive function and psychopathology: A neurodevelopmental perspective. *Annual Review of Clinical Psychology*, 16, 431–454. (Publisher: Annual Reviews). DOI: <https://doi.org/10.1146/annurev-clinpsy-072319-024242>

TO CITE THIS ARTICLE:

Paskewitz, S., Brazil, I. A., Yildirim, I., Ruiz, S., & Baskin-Sommers, A. (2024). Enhancing Within-Person Estimation of Neurocognition and the Prediction of Externalizing Behaviors in Adolescents. *Computational Psychiatry*, 8(1), pp. 119–141. DOI: <https://doi.org/10.5334/cpsy.112>

Submitted: 11 January 2024

Accepted: 26 June 2024

Published: 26 July 2024

COPYRIGHT:

© 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Computational Psychiatry is a peer-reviewed open access journal published by Ubiquity Press.