Predictive modeling of neurobehavioral state and trait variation across development

Sara Sanchez-Alonso*, Richard N. Aslin

Haskins Laboratories & Yale University, United States

ARTICLE INFO
Keywords:
State
Trait
Neurodevelopment
Variance
Functional connectivity
Mixed growth model
Language

ABSTRACT
A key goal of human neurodevelopmental research is to map neural and behavioral trajectories across both health and disease. A growing number of developmental consortia have begun to address this gap by providing open access to cross-sectional and longitudinal ‘big data’ repositories. However, it remains challenging to develop models that enable prediction of both within-subject and between-subject neurodevelopmental variation. Here, we present a conceptual and analytical perspective of two essential ingredients for mapping neurodevelopmental trajectories: state and trait components of variance. We focus on mapping variation across a range of neural and behavioral measurements and consider concurrent alterations of state and trait variation across development. We present a quantitative framework for combining both state- and trait-specific sources of neurobehavioral variation across development. Specifically, we argue that non-linear mixed growth models that leverage state and trait components of variance and consider environmental factors are necessary to comprehensively map brain-behavior relationships. We discuss this framework in the context of mapping language neurodevelopmental changes in early childhood, with an emphasis on measures of functional connectivity and their reliability for establishing robust neurobehavioral relationships. The ultimate goal is to statistically unravel developmental trajectories of neurobehavioral relationships that involve a combination of individual differences and age-related changes.

1. Introduction

A main goal of developmental cognitive neuroscience is to understand the systematic neurobehavioral changes observed across development and how they relate to environmental and experiential factors. Recent large-scale neuroimaging efforts (e.g., the Adolescent Brain Cognitive Development (Casey et al., 2018) [ABCD], the Healthy Brain Network (Alexander et al., 2017) [HBN] or the Human Connectome Project [HCP] Lifespan (Howell et al., 2019; Somerville et al., 2018) studies) have facilitated this goal by providing access to quantitative developmental datasets of an unprecedented breadth and size. Crucially, these ‘big data’ initiatives have made widely available both cross-sectional and longitudinal brain imaging measurements, along with behavioral and environmental assessments. These datasets are particularly useful for understanding variation in neural measures and their behavioral correlates, which may allow for characterization of change within an individual and between individuals. Achieving a characterization of these different sources of variance is crucial to develop predictive models of both normative and atypical development and ultimately growth curves that enable prediction of behavior from neural measurements. Importantly, growth curves of neurobehavioral relationships may in turn help to identify sensitive periods during development when potential interventions are most effective and to characterize the impact of environmental factors in predicting neurobehavioral development.

Despite the general agreement that characterizing neurobehavioral variation across development is crucial, it is still unclear how to best leverage these large-scale datasets to characterize the different sources of variation that contribute to neural and behavioral development (Becht and Mills, 2020; Telzer et al., 2018). In this perspective piece, we argue that there are at least two main components of variance that are essential for characterizing neurobehavioral trajectories and age-related changes: i) individual differences or trait-like patterns of variation and ii) within-subject differences or state-like patterns of variation. We elaborate on this argument by showing that neurodevelopmental research has much to benefit from combined quantification of these two sources of variance, their interactions and changes over time.

To illustrate our perspective, we can consider a commonly used
measure of brain function, functional connectivity (FC). FC refers to the degree of covariation between spatially distributed signals in the brain (Biswal, 2012; Snyder, 2016; Snyder and Raichle, 2012). The functional organization of the adult brain has been extensively characterized via FC with a variety of neuroimaging techniques, such as functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), electro-encephalography (EEG), magneto-encephalography (MEG), and electrocorticography (ECoG). FC patterns in the adult brain tend to be relatively stable within an individual, and similar patterns can be captured across subjects (de Souza Rodrigues et al., 2019; Demuru and Fraschini, 2020; Finn et al., 2015; Gordon et al., 2017a, b; Miranda-Dominguez et al., 2014). This within-individual stability in FC measures has been leveraged to understand individual differences or ‘trait’ variance by investigating neural patterns in single subjects and how they may relate to behavioral phenotypes (Bosi et al., 2018; Finn and Todd Constable, 2016; Friedman et al., 2019; Gratton et al., 2016; Greene et al., 2016; Nostro et al., 2018; Oswald et al., 2017; Rosenberg et al., 2016; Seitzman et al., 2019; Yoo et al., 2018).

Trait-related neural variation is often measured with resting-state FC. Resting-state FC refers to patterns of covarying spontaneous neuronal activity among a set of brain regions, which are observed even in the absence of external stimuli (Biswal, 2012; Biswal et al., 1997; Fox et al., 2005; Snyder and Raichle, 2012). These patterns were originally used to capture the statistical properties of neural activity that is spontaneously generated in the absence of a task (Biswal et al., 1995, 1997; Raichle, 2015; Snyder, 2016; Snyder and Raichle, 2012; Van den Heuvel and Hulshoff Pol, 2010). Importantly, resting-state FC and associated networks are highly stable and replicable within an individual and FC patterns during rest have been leveraged to understand individual differences across health and disease (Finn et al., 2015; Gordon et al., 2017a; Gratton et al., 2018; Greene et al., 2018; Satterthwaite et al., 2018; Seitzman et al., 2019). In parallel, studies have also shown that consistent and specific FC patterns emerge as a function of cognitive task (Cole et al., 2013, 2014, 2016; Gratton et al., 2016). For example, tasks engage a consistent whole-brain network organization different from resting-state networks and frequently result in higher connectivity among certain sub-networks. Furthermore, different tasks are associated with systematic and distinct network changes. Connectivity between language and visual networks increases during naturalistic viewing tasks (Betti et al., 2013) or between visual and dorsal attention networks during visual attention tasks (Spadone et al., 2015) relative to rest.

Task-related sources of variation, however, explain only a portion of the observed network variance, which is primarily accounted for by trait-related sources of variation (Gratton et al., 2018). Removal of resting-state network estimates from task-based estimates seems to enhance the more transient task-dependent effects (Cole et al., 2019; Gratton et al., 2018). Collectively, these findings show that FC measures may reflect stable traits of individuals, while also capturing more state-like transient patterns that are consistently associated with specific cognitive tasks.

Understanding the contribution of, and interactions between, state and trait components of variance in shaping neural and behavioral development requires a characterization of their changes over time. Characterizing these time dependent relationships across development has proven challenging, especially during periods of rapid change. During the first years of life, for example, the human brain undergoes dramatic and rapid maturational changes in both anatomy and function. This is a critical time during development in which a number of normative learning mechanisms and susceptibility to social disorders emerge (Huartseren et al., 2016; Keumen et al., 2017; Vasung et al., 2019). During this period, the reliability and stability of neural and behavioral measurements—which are crucial for identifying normative changes expected as a function of age—may be affected by state-like patterns of developmental variation that modulate concurrent trait-like variations (Blasi et al., 2014; Herting et al., 2018). Importantly, signal fluctuations at the individual-level during early childhood may be a reflection of normative changes and variable trajectories associated with typical development. Furthermore, research on developmental change indicates that signal variability is likely to increase during the acquisition of new skills (Adolph et al., 2003, 2008; McMurray, 2007). These acquisition periods are often characterized by the stochastic presence/absence of a given skill, until a stable period of daily expression is reached. Importantly, these periods of frequent regressions and sudden alterations in the rate of change may reflect developmental phases of particular susceptibility to environmental input. Achieving a comprehensive understanding of neurobehavioral variation during these periods of rapid change will likely require fine-grained neurobehavioral measures that reflect signal variability and contributions of both state- and trait-dependent components of variance.

Finally, developmentally-relevant time-dependent neurobehavioral relationships can also be observed within a single task. These are fluctuations in neural activity that change over relatively short timescales and have been primarily studied via invasive electrophysiological recordings of single cells and non-invasive electroencephalograms (Allen et al., 2018; Criado et al., 2008; Rey et al., 2015; Shou et al., 2020). These methods have a higher temporal resolution than fMRI and allow investigations into the adaptability and dynamics of brain functioning and their cognitive correlates. In contrast, studies using fMRI have often implicitly assumed that the observed spatial and temporal patterns are stationary throughout the length of the scanning session (e.g., during a task or resting-state). Nevertheless, research into the time-varying aspects of FC measures indicates that dynamic FC patterns can also be identified via fMRI. For example, there are correspondences between features of FC over short periods of time (e.g., variability or modularity shifts) and measures of task performance (e.g., response-time or accuracy) (Bassett and Mattar, 2017; Bertolero et al., 2020; Kao et al., 2020; Mattar et al., 2018). These dynamic FC patterns seem to be modulated by learning and experience, which has raised questions regarding how environmental stimuli and associated learning experiences shape the developing brain. Bassett and colleagues (Bassett et al., 2011), for example, investigated the reconfigurations of human brain networks during learning in a group of young adults and showed that learning is dependent on the flexibility of brain connections to adapt and change in relation to environmental input. Using tools from network science and graph theory, they in turn used variability in network structure to make predictions about the amount of learning in subsequent experimental sessions. Similar results have been found in studies investigating the effects of learning on brain network reconfigurations across a variety of cognitive domains such as spatial, motor, perceptual or value learning and also with respect to cognitive load (Antzoulavos and Miller, 2014; Bassett et al., 2015; Bertolero et al., 2020; Gerraty et al., 2018; Mattar et al., 2018). These findings provide evidence for the dynamic nature of FC patterns and highlight the potential of understanding how the dynamic social environment and individual-specific learning experiences contribute to shape brain networks and cognitive development.

In the remainder of this perspective piece, we posit that mapping neurodevelopmental variation requires combining, both conceptually and quantitatively, state- and trait-dependent sources of variance. In doing so, we consider how these orthogonal components of variance contribute to characterizing neural and behavioral relationships across development with the ultimate goal of enabling prediction of behavior from neural patterns. We first examine examples of state- and trait-dependent sources of variance found across anatomical (cortical folding patterns), functional (FC) and behavioral (language) measurements. These examples will illustrate how measures of neural and behavioral variation can be characterized by a combination of these two orthogonal components of variance. We then consider state and trait relationships that occur concurrently in relation to a given neural or behavioral measurement with a focus on brain FC. We discuss how state- and trait-dependent sources of variance and their interactions can be formally quantified via models that partition the variance into within- and between-subject components. Crucially, we argue that modeling
non-linear growth curves along with state-like and trait-like patterns of variation is necessary to comprehensively map neurobehavioral trajectories across development. Finally, we examine key factors that may affect the reliability of developmental measures for establishing robust neurobehavioral relationships.

2. Mapping state and trait components of neurodevelopmental variation

Neuroimaging studies have traditionally quantified data across individual subjects with the goal of drawing inferences regarding general patterns of brain activity that are common across groups of people (Becht and Mills, 2020; Gordon et al., 2017b; Madhyastha et al., 2018; Telzer et al., 2018). In other words, human brain function and behavior have been studied extensively via group-averaging, which focuses on the population mean, or by investigating how individuals vary relative to the mean. To map human neurodevelopment across age, however, it is necessary to study how a specific person differs from the group-level pattern with respect to neural and behavioral variables and how individuals themselves vary over time. This necessitates consideration of ‘within-subject’ or ‘state’ variance components, which are present alongside ‘between-subject’ or ‘trait’ variance components. Statistically, these two sources of variance are orthogonal and therefore can be characterized independently of each other.

Recent large-scale neuroimaging efforts have facilitated the process of quantifying individual variability across a population and a number of studies have shown that individuals exhibit FC patterns that differ from the group-level pattern (Gordon et al., 2017a; Marek et al., 2019; Seitzman et al., 2019). These individual differences in functional brain organization have been associated with stable, trait-like systematic variation in behavior. Crucially, it is unclear to what extent these individual-specific variations and associated network organizations are stable within a person or exhibit state-related changes. This core assumption is fundamental for neurodevelopmental research since humans undergo rapid neural and behavioral state- and trait-related change across development, especially during the early years of life. Therefore, accounting for these two sources of variance, and their interactions, is a prerequisite to comprehensively quantify differences across individuals, as well as variation within a person as a function of time-dependent genetic, neural, behavioral and environmental factors. Understanding individual differences as a source of variation in a population, however, has often ignored the orthogonal source of within-subject variance (Geerligs et al., 2015; Seghier and Price, 2018; Sharda et al., 2015; Geerligs and Tsvetanov, 2017). Here we discuss these two ways of quantifying variance, namely trait- and state-dependent variance, and how modeling them in combination is essential for advancing neurodevelopmental research.

**Trait-dependent** variance refers to a signal that exhibits variation across units of measurement relative to the group mean – in this case units would be individuals. Formally, this can be expressed with the following sum of squares (SS) equation:

\[
SS_{\text{tr}} = \sum_{i=1}^{n} (x_i - \bar{x})^2
\]

Where

- \( i \) refers to the \( i^{th} \) individual;
- \( n \) is the number of individual observations;
- \( x_i \) is the individual observations;
- \( \bar{x} \) is the individual mean

Importantly, this variance component can describe a signal that may show minimal state-dependent (e.g., stable over time) variance or that may show large state-dependent (e.g., unstable over time) variance. In other words, the ‘trait’ component of variance is orthogonal to the ‘state’ component.

In contrast, **state-dependent** variance refers to a signal that shows variation over repeated observations of the same unit (e.g., a single child), such that the unit of measurement varies as a function of measurement instance (e.g., over time or over different experimental conditions such as resting-state versus task). Formally, this can be expressed with the following equation:

\[
SS_{\text{st}} = \sum_{i=1}^{n} (x_i - \bar{x}_i)^2
\]

Where

- \( i \) refers to the \( i^{th} \) individual;
- \( n \) is the number of individual observations;
- \( x_i \) is the individual observations;
- \( \bar{x}_i \) is the individual mean

To exemplify these two sources of variance, we will use cortical folding patterns because they represent a natural biological process that exhibits both trait-dependent variance (i.e., adult cortical folding patterns are highly variable across people but relatively stable at the individual level) as well as state-dependent variance (i.e., rapid and highly dynamic folding during early neurodevelopment which also differs across children) (Fig. 1). In the adult brain, there is relatively minor within-subject (i.e., state-dependent) variance in cortical folding patterns. However, cortical folding patterns vary widely across individuals, even between pairs of monozygotic twins (Van Essen et al., 2019) (i.e., trait-dependent variance). Importantly, a number of studies have shown that cortical folding patterns, such as sulcal depth, are related to FC patterns (Mueller et al., 2013). More generally, trait-dependent variance in brain anatomy and function has been associated with behavioral variation that is stable within an individual but varies drastically across individuals (Bayly et al., 2014; Garcia et al., 2018a; Mueller et al., 2013; Satterthwaite et al., 2018; Seitzman et al., 2019). This type of variance that shows high stability at the individual level is what we refer to as trait-dependent variance.

In contrast, an example of state-dependent variance are cortical folding patterns in the third trimester of prenatal development, a period of rapid cortical expansion, during which folding patterns show high within-subject variability (Garcia et al., 2018b). Alterations in the development of cortical folding patterns during this period have important consequences for healthy development later in life and indeed have been associated with a range of cognitive and emotional disorders (Garcia et al., 2018b; Sun and Hevner, 2014). Several studies suggest that cortical folding patterns are driven by mechanical tension along long-distance axons in the white matter since the outer gray matter shows more rapid growth than the underlying white matter (Bayly et al., 2014; Garcia et al., 2018a; Kroenke and Bayly, 2018). Therefore, folding variability may impact within-subject connectivity and consequently variation in the connectivity profile of different brain areas. Despite the complex cortical folding patterns that emerge early in development and their relevance for understanding developmental disorders, to our knowledge, there are no studies that have attempted to characterize how this basic anatomical phenomenon varies as a function of both trait- and state-dependent sources of variance in the same sample of individuals.

Another instance of trait- and state-dependent sources of variance can be observed in measurements of brain FC (Fig. 1). Although FC measures exhibit both state and trait variation, studies have traditionally focused on only one of these two sources of variance in a simplifying effort to address fundamentally different questions. An example of trait-dependent variance is variability in FC patterns during resting-state, which are relatively stable within an individual. Indeed, these resting-state FC patterns can be used to ‘fingerprint’ an individual; in other words, to identify that particular individual relative to others based on their FC patterns (Finn et al., 2015; Gordon et al., 2017b;
Miranda-Domínguez et al., 2014). In contrast, FC measurements associated with a given cognitive task are an example of state-dependent variance. For example, FC patterns during movie-watching are distinct from resting-state FC patterns (Betti et al., 2013; Demirtaš et al., 2019) and indeed movie-watching and rest have been associated with brain-wide differences in FC (Sanchez-Alonso et al., 2020). There is also evidence of systematic within-subject FC variability as a function of cognitive task (Cole et al., 2013, 2014). Moreover, variations in FC can be used to decode the specific task in which an individual is engaged, thus suggesting that these state-dependent FC patterns are not random but vary systematically within a person in a way that can also be captured across the entire group (Chen et al., 2018). Collectively, these data illustrate how both trait- and state-dependent components of variance in brain FC measurements can uniquely contribute to mapping of neurobehavioral relationships. Two key questions in neurodevelopmental research are i) whether FC measurements that show trait-dependent variance also exhibit behaviorally-relevant state-dependent changes across development and ii) how these two sources of variance change over age in relation to experiential and environmental factors.

Finally, trait- and state-dependent variance is observed at the behavioral level in measures of expressive language skills (Fig. 1). In adults, expressive language skills are relatively stable within an individual and show minimal day-to-day variation (Thiessen et al., 2016). That is, our language skills do not change much during adulthood. Note, however, that specific language measurements may vary substantially among adults, for example, as a function of education level, socioeconomic status or language proficiency (e.g., in second language learners). In contrast to adulthood, language expressive skills vary greatly within an individual during early childhood, especially in the first five years of life. As we acquire language there is tremendous day-to-day, and month-to-month change in measures of language skills (Adolph et al., 2008; Feldman, 2019; McMurray, 2007; Thiessen et al., 2016). For example, whereas the first words are produced at around 12 months of age, it is only about three years later (48–60 months) that most children are able to use grammar at near-adult levels and to construct narrative discourse (Feldman, 2019). Measures of this robust state-dependent language variation in early childhood have indeed been harnessed to identify key milestones of typical language development and to study how language deficits can be indicative of neurodevelopmental disorders. Here, language illustrates a key example of why characterizing sources of state-dependent and trait-dependent variance is essential, especially for neurodevelopmental processes that exhibit critical and rapid changes during early childhood. Although these two components of variance are orthogonal, they can be observed concurrently. In other words, a given neural or behavioral measure can simultaneously exhibit within-subject and between-subject change – a point we turn to in the next section.

3. Concurrent changes of state and trait neurodevelopmental variation

Having established that neurobehavioral measures can be quantified by examining their trait- and state-dependent variance components, we now turn to the fact that they may also exhibit effects that are interactive. That is, state-dependent variance may exhibit change with age/task as a function of change in trait-dependent variance. To isolate such state-trait variance relationships, it is essential to formalize them in a hierarchical multi-level model – a framework we will return to later. Here, we focus on brain FC measurements to formulate a scheme for how to conceptualize concurrent changes in FC variability along both state-dependent and trait-dependent variance components.

To exemplify this idea, we can consider research in neurodevelopmental disorders, which has often focused on characterizing trait-dependent sources of variance. Research on autism spectrum disorder (ASD), which is characterized by impairments in social communication and restrictive, repetitive behaviors, points to distinct FC patterns that distinguish ASD from typically developing controls across rest and task (Easson et al., 2019; Hull et al., 2017; Long et al., 2016; Xu et al., 2019). Less frequent, however, are studies that characterize both task-specific (state) and disorder-specific (trait) effects within the same sample. One such study was conducted by Jasmin and colleagues (Jasmin et al., 2019), who examined FC variation in individuals with ASD and controls as they participated in two social interaction tasks that required different levels of social demands, as well as in a resting-state session. They were able to quantify both trait and state effects on FC, as well as their interaction, which revealed different levels of interregional correlation as a function of task and rest between individuals with...
ASD and typically developing controls. Specifically, ASD-specific cortico-cortical interactions showed an increase during task and a decrease during rest in comparison to controls (i.e., contextual or state-dependent differences, Fig. 2). In contrast, striato- and thalamo-cortical interactions increased across rest and tasks in comparison to controls (i.e., core or trait-dependent differences, Fig. 2). These rest-task effects and group-level differences allowed the authors to identify these striato- and thalamo-cortical connections as being core to ASD impairments. This example illustrates how leveraging state-dependent and trait-dependent sources of variance in the same groups of participants can contribute to a characterization of FC variability in relation to neurodevelopmental disorders. These state-trait variance relationships are crucial to achieve a comprehensive understanding of neurodevelopmental variation as it can reveal nuanced neural patterns that are driven by specific tasks and that may aid in characterizing population-level differences.

State-dependent neurodevelopmental variation can be observed as a function of cognitive task, but importantly, also in relation to age. For example, there is variation in FC patterns during rest as a function of an individual’s age (i.e., across years, but also months or days), which has been associated with brain maturation processes and experience (Betzel et al., 2014; Gao et al., 2015; Grayson and Fair, 2017; Hoff et al., 2013; Hutchison and Morton, 2015; Oldham and Fornito, 2019; Smyser et al., 2010). Network FC at birth, quantified using graph theoretic techniques (i.e., degree and betweenness centrality measures), is strongly determined by local anatomy (De Asis-Cruz et al., 2015; Van Den Heuvel et al., 2015), which evolves into a more distributed organization that supports the establishment of primary sensory networks within the first two years of life (Egebrecht et al., 2017; Gao et al., 2015). For a review on the development of large-scale functional networks see (Grayson and Fair, 2017). At the group level, within- and between-network FC measures are relatively stable in children and adults of the same age, thus allowing prediction of an individual’s age based on variations in FC patterns across age (Dosenbach et al., 2010). More recent studies have focused on investigating sources of state-dependent variation in FC measures and how they relate to age-specific variation. Geerligs and colleagues (Geerligs et al., 2015) investigated age- and state-specific effects on FC in a large-scale developmental sample (18–88 years). They computed FC matrices across three different brain states: a resting-state session, a movie-watching session and a sensorimotor task. Most of the variance across states was explained by commonalities in FC patterns that are found in each state (63–87%). Importantly, they also quantified the effects of age on FC across states by calculating the correlation between age and FC. They found that the percentage of FC variance explained by trait-dependent effects (FC shared across states) was approximately equal to the percentage of variance explained by state-dependent effects (state-specific FC) (Fig. 3). These findings provide evidence for the important role of both state- and trait-dependent effects in shaping FC patterns across development. Understanding these concurrent state- and trait-dependent effects is therefore crucial to obtain a comprehensive picture of neurodevelopmental change given the rapid and dramatic age-related effects observed during childhood and adolescence.

So far, we have considered studies that assume brain FC to be a stationary measure throughout the length of the scanning session. As a result, each scan (e.g., rest or task) is considered a static snapshot of the participant’s brain function as assessed by FC. In other words, each scan session characterizes a ‘state’. An alternative is to consider state and trait relationships that occur ‘within’ a specific task or scan session. For example, a growing number of studies have focused on time-varying analyses of FC, which investigate how brain function reconfigures itself at the scale of seconds to minutes (Gonzalez-Castillo et al., 2019; Saggar et al., 2018; Shine et al., 2016). In these studies, the term ‘state’ refers to whole-brain stable FC configurations recurring across participants and time (Allen et al., 2014; Damaraju et al., 2014; Gonzalez-Castillo et al., 2015; Valasina et al., 2019). Note, however, that these two definitions of ‘state’, whether it be a whole scan session of 6–10 min or FC configurations that are 30 s in duration, can be characterized by FC measures that show within-subject variability.

Time-varying FC dynamics have been modelled during movie-watching and story-listening, which show stable shifts in patterns of FC across time and participants (Betzel et al., 2019; Bolton et al., 2019; Manning et al., 2018; Simony et al., 2016). These FC shifts in turn relate to scene boundaries identified by human observers (Baldassano et al., 2017, 2018; Silva et al., 2019). Similarly, brain FC networks change dynamically as a function of cognitive demands (Antzoulouatos and Miller, 2014; Bassett and Mattar, 2017; Braun et al., 2015; Mattar et al., 2018). These studies often quantify FC changes via shifts in modularity and community allegiance using tools from network science and graph theory. State-dependent within-subject variability is observed as a function of learning or interactions with the environment, for example...
as a result of consecutive stages in a learning paradigm. Along these lines, Bassett and colleagues (Bassett et al., 2011) showed that the amount of person-specific FC flexibility, as measured by nodal flexibility, can predict the amount of learning of a specific participant in a subsequent experimental session. Importantly, network variability was not accounted for by trait-dependent sources of variance. That is, this variation is not a stable signature of an individual’s functional brain organization, but instead changes as a function of the learning experience. Collectively, studies investigating dynamic reconfigurations of brain networks within a task or movie scan may be particularly relevant for understanding developmental changes and, specifically, how children adapt their existing brain function to the contextual demands of the environment. Nevertheless, despite the growing literature on time-varying FC analyses, criticisms have emerged regarding their statistical validity and physiological origins, as well as their accuracy and reliability. Furthermore, recent studies suggest that time-varying FC during resting-state, in which there is no induced state (e.g., as a result of a task or movie), may be the result of sampling variability of static FC, head motion, and/or changes in arousal (Hindriks, 2015, Laumann, 2017) (see Lurie, 2019 for a review on main concerns and controversies regarding time-varying FC in resting-state fMRI).

4. Combined quantification of state and trait-dependent variation to characterize neurodevelopmental trajectories

Combined quantification of state- and trait-dependent sources of variance and their interactions requires models that can partition the variance into within- and between-subject components. While this is perhaps trivial, it is important to emphasize a key property of models that combine these two sources of variance: they can represent both correlated and uncorrelated (independent) variance/error. In contrast, many statistical procedures in the general linear model (GLM) family (e.g., correlation, regression, analyses of variance and factor analysis) cannot be used to partition correlated versus independent variance since they assume data independence (i.e., uncorrelated error). A common alternative is the use of mixed models, such as hierarchical linear models (HLMs), which allow partitioning of the variance into within- (correlated) and between-subjects (independent) components in the same model, and thus represent the effects of variables ‘nested’ at different levels of measurement (Duncan and Duncan, 2004; Garson, 2012; Hesser, 2015) (additional cross-sectional and longitudinal modeling approaches are discussed in Becht and Mills, 2020; King et al., 2018; Madhyastha et al., 2018; Telzer et al., 2018). HLMs are an alternative

---

**Fig. 3. Concurrent Age and State Effects.** Correlation between age and functional connectivity for each pair of regions of interest (ROI) across brain states (resting-state, movie-watching, sensorimotor). \( r \) = similarity of age effects between states, \( H_0 \) = expected value for the similarity between states if there were no true state differences (based on split-half analyses). Based on these two correlation coefficients, the following percentages were calculated: overlap = percentage of variance explained by trait effects (i.e., variance in the FC matrices that is shared across states), state.diff = percentage of variance explained by state effects (i.e., variance due to state-related changes in connectivity). Reproduced, with permission, from Geerligs et al. (2015).
HLMs are particularly useful for analyses of longitudinal data, where repeated observations of the same individual are measured over time. Longitudinal data exhibit autocorrelation since observations at two consecutive points for a given individual are likely to be more similar to one another than observations at those same time points for two different individuals. A common type of longitudinal HLM is mixed growth models, in which age is modeled as a within-subject predictor on some measurable phenomena of the individual (Duncan and Duncan, 2004; Hesser, 2015). Growth modeling allows identification of both patterns of change in developmental curves over time and the effects of different variables (e.g., sex or IQ) on the intercepts and slopes of the curves over time. Graphically, this type of analysis can be visualized with a ‘growth curve’ for each individual, in which the X-axis is age and the Y-axis is amount of variance in a given dependent variable. Fig. 4 shows examples of idealized developmental growth curves.

The growth curves illustrated in Fig. 4 vary at the individual level and can be characterized as a function of state- and trait-dependent sources of variance. These developmental patterns can be analyzed with respect to their curvature over time (e.g., speed or acceleration rates). Statistical models frequently assume that growth curves follow a linear trajectory. There are real-world examples of linear trajectories in which there is no change over time, such as gene-related variance, which is expected to stay constant across the lifespan within the individual, while at the same time displaying between-subject variability (Fig. 4A). The assumption of linear fits in developmental growth curves, however, cannot comprehensively capture developmental changes (Madhyastha et al., 2018; Telzer et al., 2018). Linear trajectories often assume constant within-subject variability and increasing between-subject variability, which would imply that a given measure never stabilizes (Fig. 4B).

Alternatively, non-linear relationships can be modelled statistically by adding non-linear terms such as a log or power function to the predictors in HLMs (Grimm et al., 2011; King et al., 2018). Fig. 4C–D show examples of non-linear fits between age and amount of signal variance for a given phenotype. A simple example of a non-linear relationship is variability in height across development, which can be characterized with a non-linear curve that shows a rapid increase and high state-dependent (within-subject) variation in early development (Black and Krishnakumar, 1999). Height stabilizes during adolescence and remains stable during adulthood, which is a period of minimal state-dependent variation, but high trait-dependent (between-subjects)
variation (Fig. 4C). Another example of a non-linear relationship observed across development is the process of synaptic pruning. Synaptic pruning increases rapidly between ages 2–10 years and peaks during adolescence (Juraska and Willing, 2017; Petanjek et al., 2011). The total number of synapses begins to stabilize after this period and remains constant into early adulthood, when it plateaus and shows minimal state-dependent variation (Fig. 4D). These state and trait dynamics show how non-linear trajectories can capture state-to-trait transitions across development. They also illustrate the necessity of considering temporal dynamics of state-trait variation and non-linear growth models to characterize developmental trajectories.

Characterizing concurrent effects of state-trait relationships can be particularly relevant for understanding neurobehavioral variation during periods of rapid developmental change or greater sensitivity to learning (Adolph and Robinson, 2011). These are periods of increased variability that may be particularly useful to delineate the path of change for a given neurobehavioral relationship (Badde et al., 2020; Bailey et al., 2001; Knudsen, 2004; Newport et al., 2001; Sourav et al., 2019). Consider language development during the first years of life, a time characterized by multiple sensitive periods across a series of domains (e.g., tuning of native phonetic categories, acquisition of grammar) (see Glossary) (Knudsen, 2004; Nelson and Gabard-Durnam, 2020; Newport et al., 2001; Takanishi and Hensch, 2013; Werker and Hensch, 2015). These periods of rapid developmental change are associated with increased brain plasticity for encoding specific environmental stimuli through experience (Hensch, 2005; Knudsen, 2004; Sengpiel, 2007; Werker and Hensch, 2015). Of relevance, sensitive periods are characterized by rapid changes in the individual, which may be represented as growth-curves with high within- and between-subject variability at the neural and behavioral levels (Fig. 4E), which over time stabilize (Fig. 4F) (Gabard-Durnam and McLaughlin, 2019; Hensch, 2005; Knudsen, 2004). During these periods, exposure to specific stimuli and experiences is necessary for cortical specialization and development of neural circuits. For example, in the case of language development, the absence or limited exposure to language during the early years of life can have long-lasting negative consequences (Newport et al., 2001; Sakai, 2005; Werker and Hensch, 2015). These are therefore periods of high vulnerability to adverse events in which the timing and exposure to specific types of experiences is essential for healthy development of neural circuits.

Quantifying both state- and trait-sources of neurobehavioral variation in relation to sensitive periods is crucial to comprehensively understand the impact of environmental stimuli and the effects of learning in typical and atypical development. To illustrate this point, we can consider changes in language neurodevelopment. The first five years of life are characterized by sensitive periods of high within- and between-subject variability in language-related measures as the child achieves key developmental milestones (Fig. 5). During this period, variability in the acquisition of language skills may show accelerating or decelerating rates of change, for example as children acquire new words (McMurray, 2007). Furthermore, the acquisition of specific skills may be expressed intermittently, showing a period of variability during which the presence and absence of a given skill appears to be stochastic until it stabilizes (Adolph et al., 2003, 2008; Adolph and Robinson, 2011). These periods of regressions and sudden transitions in the rate of change are a reflection of the learning process and contribute to the variation observed during development as the child achieves main language milestones. In parallel to variation at the behavioral level, the brain undergoes rapid anatomical and functional changes such as the development of cortical folding patterns, synaptic pruning and changes in neural circuits (Cusack et al., 2018; DiMartino et al., 2014; Grayson and Fair, 2017; Keunen et al., 2017). These sensitive periods of language neurodevelopment, characterized by high within- and between-subject variability at the neural and behavioral levels, may be the most essential periods for discrimination of state and trait variance. That is, these may be key neurodevelopmental time windows in which children exhibit maximal change in state and trait variance. Importantly, these are also periods in which the child is most vulnerable to environmental input and adversity (Gabard-Durnam and McLaughlin, 2019; Nelson and Gabard-Durnam, 2020).

This quantitative framework fundamentally necessitates large-scale datasets that have the statistical power to capture neural and behavioral variability at the individual level across age. Furthermore, statistical analyses based on multivariate non-linear growth models that account for state- and trait-dependent sources of variance at both neural and behavioral levels are a prerequisite to fully understand neurodevelopmental variation. In combination with methodological and analytical advances, this framework may ultimately provide a comprehensive characterization of neurodevelopmental trajectories and periods of heightened plasticity in which the brain may be particularly sensitive to environmental stimuli and clinical interventions.

**Fig. 5. Mapping State and Trait Dynamics Across Sensitive Periods of Language Neurodevelopment.** Two sensitive periods are highlighted in blue and red, which span different developmental windows. Each period is marked by distinct transitions in state and trait variation, which are highlighted as the most saturated portions under the curve and highlighted with a vertical dotted arrow. These hypothetical maximal state and trait variance epochs represent a phase during language neurodevelopment, which may demarcate key neural and behavioral change, both within and across children. These windows of maximal neurobehavioral variation require quantification of concurrent state and trait variance dynamics in order to map personalized neurodevelopmental language trajectories, as well as maximal vulnerability to disease both within and across children.
5. Reliability and reproducibility of state and trait measurements

A critical consideration when attempting to map neurobehavioral state- and trait-dependent variation is the reliability of the measurement, which impacts the reproducibility of the overall effect. Fundamentally, failures to obtain reproducibility of the effect may reflect lack of statistical power and consequently the need for larger samples, which may be addressed via large-scale consortium-level datasets (Button et al., 2013; Hedge et al., 2018; Marek et al., 2020; Streiner et al., 2015). Large-scale datasets crucially contribute to avoiding overfitting, which is a common concern in small samples due to the larger variability in estimates, which may lead to mistakenly fit sample-specific noise as true signal (Gelman, 2010; Ioannidis, 2008; Yarkoni and Westfall, 2017). Large samples notwithstanding, obtaining replicable neural and/or behavioral effects in developmental studies also requires careful consideration regarding the reliability of the measure. For instance, how consistent is a given measurement across repeated tests and/or sessions over time? This is particularly important if the ultimate goal is to develop predictive models of normative development at the individual-subject level (Herting et al., 2018). Furthermore, clinical applications for atypical development necessitate high reliability for obtaining actionable measurements at the individual level (Streiner et al., 2015). Here, we articulate some key general considerations that will impact reliability: i) sampling amount (i.e., the amount of data necessary to avoid measurement error for a given measure), ii) repeated sampling over time (i.e., how to separate within-subject time-dependent sampling error from true signal reflecting within-subject change), and iii) choice of dependent variable(s) and experimental paradigm.

The reliability of fMRI measurements is strongly dependent on scan length and longer scanning times are often required for individual-level estimates relative to group-mean estimates. Most recent fMRI studies investigating the amount of data required to avoid measurement error have focused on resting-state FC, primarily in adult populations. Early studies on this topic showed high to modest reliability of group-mean FC estimates (Dijk et al., 2011; Shehzad et al., 2009). With respect to single-subject functional connections, reliability has often been statistically quantified with the intraclass correlation coefficient (ICC), which measures the proportion of the total variability in a measure that is due to true between-subject variability (Caceres et al., 2009; Fournier et al., 2014; Liljequist et al., 2019). fMRI estimates have only modest individual-level reliability with a 5-min resting-state scan (Anderson et al., 2011) and individual-level reliability greatly improves when the scan length increases from 5-min up to 13-min, especially for scans obtained during the same session (Birn et al., 2013; Noble et al., 2017). Other studies suggest that scanning times longer than 25-min may be needed to reliably identify single-subject resting-state FC patterns across sessions collected within a person (Anderson et al., 2011). Similarly, reliability of FC measures during task-based paradigms is dependent on scan length, as individual differences between subjects become more reliable with longer scanning time (Gordon et al., 2017b; Shah et al., 2016). Importantly, these scan time requirements are challenging to meet in large-scale developmental initiatives because babies and young children are often unable to remain awake and motionless during long resting-state sessions (Sanchez-Alonso et al., 2020; Vanderwal et al., 2019). Consequently, large-scale consortium-level initiatives often collect data during 5–10 min resting-state sessions, which often makes it challenging to robustly single-subject predictive models.

Another relevant consideration is how repeated sampling over time points affects reliability of the measurement. When testing a subject at two different points in time, it is crucial to separate within-subject time-dependent sampling error and additional spurious sources of within-subject biological variability (e.g., diurnal rhythms or metabolic state) from true within-subject change as a result of experience or brain maturation (Hodkinson et al., 2014; Laumann et al., 2015; Shannon et al., 2013). Repeated sampling is often necessary in order to distinguish meaningful state variation from noise, especially during developmental periods of rapid change in which larger within-subject variability is expected (Adolph et al., 2003, 2008; Adolph and Robinson, 2011). Current large-scale datasets, however, often rely on single-session data acquisitions which affects measurement reliability and the ability to establish robust neurobehavioral relationships (Adolph and Robinson, 2011; Herting et al., 2018; Rush and Hofer, 2017).

Finally, the reliability of the measurement is also impacted by the type of dependent variable (e.g., anatomical versus functional MRI) and experimental paradigm (e.g., task versus rest). While anatomical studies often achieve clinically-acceptable intra- and inter-session reliability, such standards are harder to obtain with fMRI estimates because they are often subject to larger within- and between-subject variability (Bennett and Miller, 2010; Madan and Kensinger, 2017; Zuo et al., 2014, 2019). Regarding the choice of experimental paradigm in fMRI studies, eyes-open resting-state sessions often yield stronger correlations relative to eyes-closed sessions (Dijk et al., 2011; Noble et al., 2017). Studies on the reliability of FC measurements obtained across different cognitive states beyond rest is limited and the focus has been primarily on data obtained via task-based paradigms. Crucially, the unconstrained nature of the resting-state, in which a variety of brain states are sampled, may necessitate large amounts of data, but it is unknown whether such timing requirements also apply to data acquisitions that are more constrained and focused on more homogeneous task-based brain states (Elliot et al., 2020; Shah et al., 2016; Vanderwal et al., 2019).

A complementary method to traditional task-based designs are naturalistic paradigms, which sample a more constrained set of states than resting-state paradigms and are particularly well-suited for FC analyses and developmental studies (Cantlon, 2020; Emerson et al., 2015; Vanderwal et al., 2019). Naturalistic methods can take a number of forms and they usually consist of stimuli that require rapid integration of real-time information (Bottenhorn et al., 2019; Sonkusare et al., 2019). Examples of naturalistic methods include paradigms aimed at investigating neural and/or behavioral activity as the participant is exposed to natural scenes, is required to reason about problems, listens to music, or interacts with familiar individuals (e.g., the mother or father). One of the most commonly implemented types of naturalistic paradigms has, so far, been movie-watching. Reliability of FC measurements during movie-watching paradigms seems to be slightly better than during rest (Anderson et al., 2011; Wang et al., 2017) and FC measurements during movie-watching have better individual-level predictive value in young children relative to resting-state acquisitions (Sanchez-Alonso et al., 2020). Importantly, a longer movie-watching session does not necessarily yield better single-subject predictions than a shorter one, which seems to suggest that sampling more homogeneous cognitive states may require shorter acquisition times relative to resting-state sessions (Finn and Bandettini, 2020; Sanchez-Alonso et al., 2020). Current large-scale neurodevelopmental studies leverage a range of experimental paradigms, from traditional block-designs to movie-watching and resting-state paradigms, which have facilitated investigation into the reliability of FC measures across paradigms. More research is needed, however, to understand the reliability and reproducibility of FC measurements that sample more constrained states beyond rest and their utility for investigating changes in brain function as a result of cognitive performance.

Collectively, these findings indicate that accurate characterization of brain-behavior relationships in developmental samples requires taking into account the reliability of the measurement. Specifically, it is crucial to consider measurement error and within-subject time-dependent sources of variation in order to obtain a reliable characterization of within-subject change that is the result of brain maturation and/or experience. Current large-scale neurodevelopmental datasets offer a unique opportunity to investigate state and trait variation and identify potential gaps in our understanding of neurodevelopmental change. These initiatives may open the door to future studies that address
concerns regarding measurement replicability within and across sessions, including amount of data and dense longitudinal sampling. Furthermore, as the field of cognitive developmental neuroscience moves forward, we believe it is necessary to focus on understanding both state and trait variation at the neural and behavioral level. This in turn requires consideration of how state and trait variation are measured across development, including amount of signal, repeated sampling over time, type of experimental paradigm, and choice of dependent measure—all crucial factors that impact measurement reliability.

6. Concluding remarks

Human neurodevelopmental research aims to map neural and behavioral trajectories over time across both health and disease. This goal necessitates large-scale developmental datasets that allow for capturing relevant neurobehavioral variation. Recent ‘big data’ cross-sectional and longitudinal repositories have provided access to large-scale quantitative developmental datasets with rich neurobehavioral measures. These datasets can be leveraged to characterize variation at the individual level with the ultimate goal of developing predictive models of both normative and atypical development. Several studies have already convincingly shown that individuals exhibit variation in neural and behavioral measurements that are distinct from the group-level pattern. Neurobehavioral mapping, however, tends to focus only on characterizing either state- or trait-dependent components of variance, separately from each other.

In this perspective piece, we have argued that combined quantification of both within-subject (state) and between-subject (trait) variation across neural and behavioral measurements is essential to comprehensively characterize developmental change and age-related trajectories. We have described how these two sources of variance can be modelled in the same quantitative framework via non-linear mixed growth modeling. This is particularly important given that changes of state and trait neurobehavioral variation may occur concurrently across development. Therefore, person-specific neurodevelopmental variation may be related to population-level variation and, crucially, these relationships can be leveraged in the context of predictive-modeling approaches (Becht and Mills, 2020; Rosenberg et al., 2018; Telzer et al., 2018; Varoquaux and Poldrack, 2019). Finally, capturing state-trait dynamics and interactions may be particularly important to characterize sensitive periods of brain development—periods of heightened plasticity characterized by experience-dependent changes in brain function—in which there is high within- and between-subject variability.

7. Outstanding questions

- Building on current large-scale developmental datasets (e.g., ABCD, HCP Lifespan, HBN), what type of data (i.e., age range, sampling rate, amount of data, neurobehavioral measurements) are needed to comprehensively characterize neurobehavioral trajectories across development?
- In considering neurodevelopmental change, the question arises as to the experimental paradigms best suited to capture state and trait dynamics across development—are naturalistic paradigms (e.g., movie-watching or story-listening) better suited than highly controlled experiments? Would a combination of laboratory-controlled experiments and naturalistic methods be able to provide a more comprehensive understanding of the developing brain?
- What types of quantitative and modelling techniques are necessary to capture state-to-trait variance dynamics in neurodevelopment beyond non-linear mixed growth models?
- How can we leverage state-trait dynamics to predict neurobehavioral variation at a later point in life?
- Can we harness multi-modal imaging methodologies (e.g., fMRI, fNIRS, EEG) to better characterize age-related trends in brain organization across development?
- How do state- and trait-dependent sources of neurodevelopmental variance change over age in relation to experiential and environmental factors?

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgment

Research reported in this publication was supported by National Institutes of Health Grant HD-037082 (R.N.A).

References

Glossary

Resting-state Networks: Pattern of covarying spontaneous neuronal fluctuating activity among a set of brain regions, which is observed even in the absence of external stimuli.

Graph Theory: Branch of mathematics focused on the study of networks or graphs, which are defined as systems of elements (i.e., nodes, vertices) and their pairwise associations (i.e., edges, connections).

State: Measurable signal that exhibits variation over repeated observations of the same unit (e.g., a single child).

State-dependent variation: Signal variation over repeated observations of the same unit (e.g., a single child), such that the unit of measurement varies as a function of measurement instance (e.g., over time or over different experimental conditions such as resting-state versus task).

Trait: Measurable signal that exhibits variation over different units (e.g., multiple children).

Trait-dependent variation: Signal variation observed across different units of measurement (e.g., multiple children).

State-Trait Interactions: Combined (i.e., interactive) signal variance along both state-dependent and trait-dependent variance components for a given measurable unit. These effects can be observed concurrently or at different points in time across development.

State-Trait Dynamics: Change in signal variance along state-dependent and/or trait-dependent variance components for a given measurable unit over time.

Time-varying (Dynamic) FC: Covariation between spatially distributed signals in the brain that capture temporal dependencies (e.g., on the order of seconds to minutes).

Mixed Growth Model: A statistical framework for modeling longitudinal data, in which repeated observations of the same individual are measured over time.

Sensitive Period: Window of time characterized by increased brain plasticity for encoding environmental stimuli through experience, during which exposure to specific stimuli is essential for healthy development (e.g., vision or language). The absence of environmental stimuli during this time can have long-lasting consequences on behavior and may lead to irreversible changes in brain function with minimal recovery. Nevertheless, plasticity for specific neural circuits can exist beyond the sensitive period such that circuits can be reshaped later in life (e.g., as a result of new learning experiences).