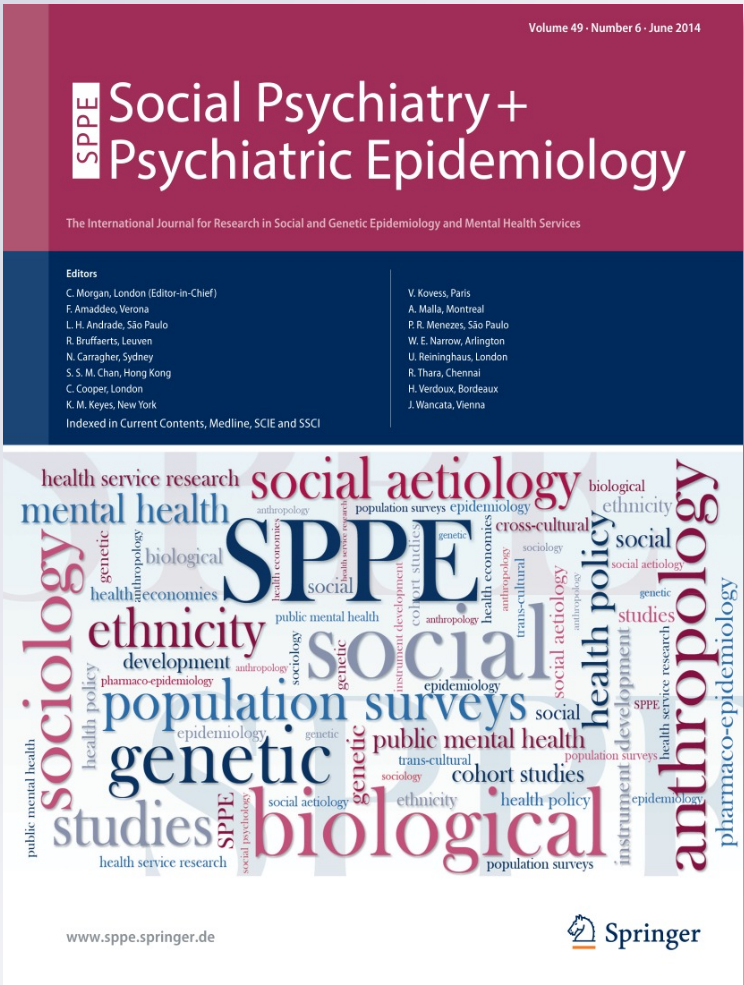
# *Translating multilevel theory into multilevel research: challenges and opportunities for understanding the social determinants of psychiatric disorders*

**Erin C. Dunn, Katherine E. Masyn, Monica Yudron, Stephanie M. Jones & S. V. Subramanian**

**Social Psychiatry and Psychiatric Epidemiology**
The International Journal for Research in Social and Genetic Epidemiology and Mental Health Services

ISSN 0933-7954
Volume 49
Number 6

Soc Psychiatry Psychiatr Epidemiol
(2014) 49:859-872
DOI 10.1007/s00127-013-0809-5




Springer

Soc Psychiatry Psychiatr Epidemiol (2014) 49:859–872
DOI 10.1007/s00127-013-0809-5

ORIGINAL PAPER

# Translating multilevel theory into multilevel research: challenges and opportunities for understanding the social determinants of psychiatric disorders

**Erin C. Dunn · Katherine E. Masyn · Monica Yudron · Stephanie M. Jones · S. V. Subramanian**

Received: 28 October 2012 / Accepted: 16 December 2013 / Published online: 28 January 2014


**Abstract**

*Introduction* The observation that features of the social environment, including family, school, and neighborhood characteristics, are associated with individual-level outcomes has spurred the development of dozens of multilevel or ecological theoretical frameworks in epidemiology, public health, psychology, and sociology, among other disciplines. Despite the widespread use of such theories in etiological, intervention, and policy studies, challenges remain in bridging multilevel theory and empirical research.

*Methods* This paper set out to synthesize these challenges and provide specific examples of methodological and analytical strategies researchers are using to gain a more nuanced understanding of the social determinants of psychiatric disorders, with a focus on children's mental health. To accomplish this goal, we begin by describing multilevel theories, defining their core elements, and discussing what these theories suggest is needed in empirical work. In the second part, we outline the main challenges researchers face in translating multilevel theory into research. These challenges are presented for each stage of the research process. In the third section, we describe two methods being used as alternatives to traditional multilevel modeling techniques to better bridge multilevel theory and multilevel research. These are (1) multilevel factor analysis and multilevel structural equation modeling; and (2) dynamic systems approaches.

*Conclusions* Through its review of multilevel theory, assessment of existing strategies, and examination of emerging methodologies, this paper offers a framework to evaluate and guide empirical studies on the social determinants of child psychiatric disorders as well as health across the life course.

**Keywords** Multilevel · Social determinants · Social and physical environments · Ecological · Context · Composition · Psychiatric disorders

E. C. Dunn (✉)
Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA 02144, USA
e-mail: erindunn@pngu.mgh.harvard.edu

K. E. Masyn · M. Yudron · S. M. Jones
Harvard Graduate School of Education, Harvard University, Cambridge, MA, USA

S. V. Subramanian
Department of Social and Behavioral Sciences, Harvard School of Public Health, Boston, MA, USA

## Introduction

The notion that health, disease, behavior, and development, including risk for psychiatric disease, is a multilevel phenomenon—or is influenced by and occurs within multiple social and physical contexts—has existed for centuries. As early as 400 BC, Hippocrates linked environmental conditions to the body's four basic substances, or humors, and described how these environmental factors could cause the humors to become imbalanced, resulting in disease [1]. Recognition of the association between features of the social and physical environment and individual-level outcomes was also reflected in other early writings, especially those that described health disparities. For example, during the nineteenth century, physicians and social reformers across Europe documented the ways living and working conditions, including child labor, were related to high rates of disease, particularly among the poor [2–4].

The observation that features of the social environment are associated with health and illness has spurred the development of dozens of multilevel theoretical frameworks. In social epidemiology and public health, this includes frameworks that describe the distribution and determinants of individual and population health and health inequalities, such as psychosocial theory [5], social production of disease/political economy of health [6], ecosocial theory [7], and social-ecological models [8, 9]. In psychology, this includes behavioral theories, such as Bandura's social cognitive theory [10]; developmental dynamic systems theories, which are particularly focused on children, including Sameroff's transactional [11] and unified theories of development [12]; and Bronfenbrenner's ecological and more recent bio-ecological theory [13, 14]. It also includes community- or setting-level theories, which focus primarily at the level of organizations [15].

Although multilevel theoretical frameworks are commonly described as guiding etiological, intervention, and policy studies, there remains an incomplete translation of multilevel theory into empirical research. That is, there is a gap in knowledge and empirical practice regarding how to best empirically test multilevel theories or translate multilevel theory into appropriate statistical models that use appropriate measures and variables. This theory-to-methods translational gap stems from the fact that multilevel theories typically provide broad models to guide thinking, without offering practical guidance on how to test the theory or portions of it. While this theory-to-methods translational gap has persisted due to the lack of explicit attention in the literature to how theories can be tested, the consequences of this gap are enormous. First, it precludes the effective testing and refinement of theory necessary for describing relationships between individuals and the contexts in which they are embedded, and prevents deeper insights into the ways both individual and environmental factors predict mental health outcomes. For example, without accurately incorporating characteristics of a individual's ecology into our statistical models, we may bias our analyses in ways that underestimate the role of social factors and overestimate the effect of individual-level characteristics (e.g., race/ethnicity, gender); this misunderstanding could not only lead to inaccurate findings regarding the importance of social determinants, but it could also result in the misspecification of intervention programs aimed at improving mental health. Moreover, this theory-to-methods translational gap also impairs our ability to conceptualize settings and ecologies. This, in turn, inhibits the development of new theories as well as theory- and research-driven interventions and policies to address risk or promote resilience and understand the level (e.g., the individual or environmental level) at which an intervention or policy may be most effective. Therefore, efforts are needed to better translate theory into empirical research and close this theory-to-methods translational gap, particularly with respect to testing multilevel hypotheses that can increase our understanding of the social determinants of psychiatric disorders.

In this paper, we outline these theory-to-method translational challenges and describe specific examples of strategies researchers are using to gain a deeper understanding of the role of environments on mental health and risk for mental illness. Although these issues are pertinent to individuals at any age, we focus on children and adolescents, given the overall dearth of research on social determinants of psychiatric disorders in youth relative to adults. We begin by defining multilevel theories and describing their core elements. Next, we inventory the challenges researchers face in translating multilevel theory into research. These challenges are presented for each stage of the research process. Finally, we review two emerging methods being used, as alternatives to traditional multilevel modeling techniques, to develop (theory generating) and test (theory testing) relationships between multilevel features and individual health outcomes. These two methods are (1) multilevel factor analysis (MLFA) and multilevel structural equation models (ML-SEM); and (2) dynamic systems approaches. Throughout, we bring together different strands of literature in an effort to raise awareness, particularly among applied multilevel researchers, of the key ingredients of multilevel theories, the strategies that have been used to study the social determinants of child mental health, and the kinds of methods needed to strengthen and expand existing knowledge. In so doing, we hope to tie together these concepts in ways that bring about a shared language among researchers from multiple disciplines that study the social determinants of mental health. Without this shared understanding, little progress can be made in generating new knowledge to guide interventions and policies. Through the practical advice and recommendations provided, we also hope this paper provides initial steps to guide thinking about ways to better bridge multilevel theories and multilevel research and ultimately bring about a new generation of multilevel studies on the social determinants of mental health. To make this article accessible to readers from multiple disciplines, we provide a brief glossary to elucidate and distinguish the terms we use throughout the manuscript (see Table 1).

## Core elements of multilevel theories

Multilevel theories explain individual or group processes and outcomes in terms of the multiple contexts in which these experiences occur. Conceptually, these frameworks examine one or more systems or environments (e.g., peer groups, schools, neighborhoods) that are most often

**Table 1** Descriptions of the core terms and concepts described throughout the paper

| Term | Definition |
|---|---|
| Theory | A theory, or conceptual framework, provides researchers an organizing framework from which to generate research questions, develop testable hypotheses, and both understand and reduce urgent health issues |
| Multilevel theory | Multilevel theories explain individual or group processes and outcomes in terms of the multiple contexts in which these experiences occur. Conceptually, these frameworks examine one or more systems or environments (e.g., peer groups, schools, neighborhoods) that are most often hierarchically nested within one another (e.g., peer groups nested within schools), and which may vary across time (e.g., developmental, historical, or intergenerational time) |
| Environment | Environment refers to the social and physical characteristics of systems, settings, or contexts relevant for children's mental health |
| Measures | Measures refer to the instruments or tools used to collect information on a given variable |
| Models | Models refer to statistical models that are used to understand relationships between two or more variables |
| Multilevel model | Multilevel models refer to any type of statistical model where the data are clustered. Clustering can refer to observations across multiple dimensions: time (e.g., repeated measures), space (e.g., clustering observations within the same environment) |
| Translating multilevel theory into empirical research | The ability to translate multilevel theory into empirical research refers to how theories of multilevel phenomenon can be tested in statistical models. Translation of multilevel theory into empirical research can be in one of two forms: conduct tests of association that are hypothesis generating (i.e., theory development) or based on hypothesis testing (i.e., theory testing) |

hierarchically nested within one another (e.g., peer groups nested within schools), and which may vary across time (e.g., developmental, historical, or intergenerational time). When studied empirically, researchers substitute the term "level" (as in group-level phenomenon) in place of system or environment. Thus, in an empirical study, multilevel can be defined in terms of the clustering of observations across time (e.g., repeated measures), or space (e.g., clustering of observations within the same environment). Here, we use the term "environment" to refer to the social and physical characteristics of systems, settings, or contexts relevant for children's mental health.

Although differing in their specific focus with respect to the populations highlighted or the constructs considered, multilevel theories all emphasize two core elements, each of which are critical to consider in effectively translating multilevel theory into multilevel research and empirically testing associations between social environmental influences and child mental health. These core elements refer to the notion that (1) there is child-level and environmental-level variability; and (2) the interplay between the child and the environment is dynamic. These two elements are elaborated below.

## Child-level and environmental-level variability

The first element multilevel theories emphasize is that there is variability, existing at the level of individuals as well as within the environments in which individuals are embedded. Thus, both individuals and their environments can vary according to one or more dimensions; individuals within environments and environments embedded in larger contexts can vary according to one or more dimensions. Some variables can also vary within individuals or environments over time.

In the next section, we address the methodological challenges that arise from investigating these three types of variability: (1) variability across children within the same environment, (2) variability within the same child across time, and (3) variability within the same child across environments. Specifically, we describe the different sources and origins of this variability and how this variation has been both theoretically understood and empirically studied. We also outline the implications of these different types of variation for conducting research.

### *Variability across children within the same environment*

No two children within the same environment are alike. In other words, if we were to look at one point in time among children in the same environment, we would expect to observe variability in children's outcomes and characteristics. Thus, there is *variation in measured characteristics*. As described by theorists, this type of variation could arise from several sources.

First, different perceptions children have about their environment may give rise to individual differences, particularly when self-reported measures are used. To that end, Bronfenbrenner and Morris [14] note

> "An early critical element in the definition of [our] bioecological model is experience, which indicates

> that the scientifically relevant features of an environment for human development not only include its objective properties but also the way in which the properties are subjectively experienced by the person living in that environment…very few of the external influences significantly affecting human behavior and development can be described solely in objective physical conditions and events.''(pp 796–797)

Thus, researchers may find that measures of the objective and subjective environment may not be as highly correlated as one would imagine due to discrepancies between what is "real" and what is perceived. This may be true for many constructs, including socioeconomic status [16].

Second, environmental effects may not be uniform across all children, but rather children may respond differently to the same environmental conditions. In other words, average effects (e.g., means or beta coefficients) may differ systematically for certain groups of children, based on the child's sex, race/ethnicity, temperament, or biological make-up. Thus, there are differences in the association between a measured characteristic and a given mental health outcome or *variation in association effects*. Theorists broadly refer to these individual features as "nature" in reference to the long-standing nature versus nurture debate, which underpins many multilevel theories. Recent empirical work on gene–environment interplay [17, 18], stress or emotional reactivity [19, 20], and resilience [21] illustrates this concept, as this research examines the ways children are differentially sensitive or vulnerable to environmental conditions and why there are individual differences in response to adversity.

Similarly, it is often presumed that an environment is the same or constant across all children. However, children may evoke different reactions or receive different treatment within the same environment, in essence experiencing a different environment, leading to individual variability. In other words, nature and nurture are inextricably linked, as described by Sameroff [12]:

> …Nature and nurture represent a unity of opposites such that neither can ever get it right on its own. Because of their interpenetration advances in our understanding of nature illuminate nurture and changes in our understanding of nature illuminate nurture (pp 11–12).

Research in the field of behavior genetics exemplifies the interrelationships between nature and nurture. For example, twin studies have shown that monozygotic twins, who share all of their genes, can have different outcomes as a result of non-shared environmental experiences, even when they are reared by the same parents [22]. It has also been shown that children create their own environments, a concept known as evocative gene–environment correlation [23, 24]. Evocative gene–environment correlation states that an individual's genetic make-up influences the responses they receive from the environment; in other words, individuals, through the effects of their genes, can affect the behavior of others. For example, some authors have argued that children who are attentive and interested in learning, as a result of some type of genetic predisposition, may be more likely to have favorable experiences with their teachers when compared with students who are distracted and disengaged [25].

There are several important research implications of this type of variability. While not always explicitly stated, these types of multilevel theories suggest that analytic methods are needed that allow for examination of between-child differences (or variation in measured characteristics) and differences in the effect of environments on children (or variation in association effects). This includes biological, psychological, and social factors as well as child-level evoked differences in the environment.

### *Variability within the same child across time*

Developmental changes are at the root of the variability observed in any one child over time. Several theories describe this type of variability. Multilevel theories focus on "development in context" and changes in children both across time and settings. For example, in his Unified Theory of Development, Sameroff [12] emphasizes the interplay between individuals and their social and physical environment over time. This transactional model of development has been further developed into the concept of developmental cascades, which emphasizes the interrelatedness of competencies across skill type within the same individual and explicitly includes biological inputs in the study of individual variation over time [26]. Multilevel theories differ from stage-based developmental theories (e.g., Piaget, Erikson) in that the latter focus on variability across children over time within a single environment and often do not consider environmental influences. Recent examples of empirical work have focused on describing variation within the same child over time and ways exposure to different types of social environments, particularly during sensitive periods of development, can adjust trajectories of mental health and other health outcomes and increase vulnerability to disease across the lifespan [27, 28]. Thus, the notion of variability within the same child across time underscores the need for longitudinal models that allow for prospectively examining changes in development over time.

*Variability within the same child across environments*

Children are embedded in multiple environments, each of which may differentially affect their mental health. The idea that individual outcomes vary based on environmental conditions and that healthier environments produce healthier individuals is at the core of several fields, especially social epidemiology [29]. The meaning of environment and the level at which the environment is conceptualized in theoretical work and studied empirically differ across disciplines, ranging from exchanges between groups of people (typically seen in sociology) to state and national policies (health policy). For example, psychosocial theories [5] focus on ways features, including dominance hierarchies, social disorganization, social isolation, and social support, predispose individuals to disease (by affecting resistance to illness), and result in racial/ethnic, socioeconomic, and other social gradients in mental health. For children, this includes empirical work documenting the salutary effects of social support on many mental health outcomes, including depression [30]. It also includes research on gene–environment interaction, which has recently expanded to include the concept of differential susceptibility to the environment [31, 32], which essentially states that some children, on the basis of a set of "sensitivity" genes, will have very different outcomes depending on whether they are embedded in an adverse or protective environment. Conversely, theories such as Doyal's [6] social production of disease/political economy of health are concerned with broader economic and political determinants of health and disease, including capitalism, and ways structural barriers prohibit individuals from living healthier lives. With respect to child mental health, this includes studies on the impact of the recent global recession and state-level mental health care expenditures [33]. The mechanisms, or pathways through which the environment influences child mental health, are varied and to date have been understudied. For example, Bandura argued that children can take cues from their environments and acquire new behaviors or adapt existing ones through watching the actions of others, a concept referred to as "observational learning" [10]. In sociology, the concept of observational learning has been adapted to focus on specific types of neighborhood and community-level social disorganization, including the theory of "broken windows". Broken windows theory suggests that failure to keep urban environments in a well-ordered condition can increase vandalism and result in elevated rates of crime [34]. Taken together, these theory elements underscore the need for analytic models that take into account between-environment differences (or variation in measured characteristics at the environment level) and provide an opportunity to examine the multiple pathways linking environmental conditions to individual mental health.

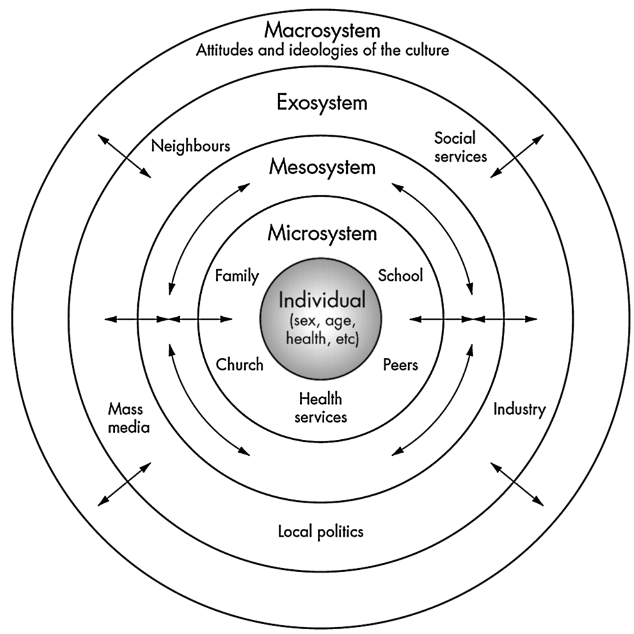



**Fig. 1** Bronfenbrenner ecological system model. This image appeared in the article: McLaren L and Hawe P (2005). Ecological perspectives in health research. Journal of Epidemiology and Community Health, *59*, 6–14. The figure is available for download at http://jech.bmj.com/content/59/1/6/F2.large.jpg

*The interplay between the child and the environment is dynamic*

The second element multilevel theories share is an emphasis on the dynamic interplay between the individual and the environment. That is, the direction of effects is not unidirectional, solely operating from the environment to the child, but rather children can affect and construct their environment. In describing social ecological models, Stokols [8] discussed this concept:

> "People-environment transactions are characterized by cycles of mutual influence, in which the physical and social features of settings directly influence occupants' health and, concurrently, the participants in settings modify the healthfulness of their surroundings through their individual and collective actions." (page 286)

Bandura coined the phrase "reciprocal determinism" to capture this circularity, arguing that human functioning is the result of an interplay between behavior, cognition, personal features, and environmental events, which all serve as determinants of one another.

This concept has been frequently depicted using Bronfenbrenner's ecological theory of human development

(Fig. 1). In this theory, Bronfenbrenner adopts a broad conceptualization of multiple nested environments and emphasizes the "reciprocal causation" between individuals and their social (as opposed to physical) environment. Although it is highly challenging to identify, measure, and model such feedback loops in an empirical study, particularly given issues of temporality, researchers will need analytic models that allow more of this type of circularity to be measured and modeled in order to effectively translate the concept of reciprocal determinism.

In summary, it is through these two core elements that multilevel-oriented theories differ from single-level theories. Specifically, multilevel theories state the need for methodological approaches and analytic methods that: (1) examine variability at the level of children and their environment both in measured characteristics and in association effects; (2) study the effect of different time scales (e.g., change in development over time; effect of specific periods of development); (3) allow for mediation pathways (or allow one to examine how changes in one variable lead to changes in another variable, which in turn predict a subsequent outcome); and (4) assess ways in which child-level outcomes are influenced by environmental-level effects and how the environment, especially the social environment, can be shaped by child-level effects. These are the key issues that must be considered in translating multilevel theories into empirical research.

## Challenges of and progress in translating multilevel theory into empirical research

Despite the existence of several well-developed multilevel theories, researchers are often hindered in their ability to effectively translate multilevel theories into empirical research due to the lack of methods and clear guidance on how to do so. As a result, the core elements of multilevel theories have not yet been fully evaluated empirically. In this section, we describe some of the major challenges and advancements that have been made in studying the core elements of multilevel theory and translating multilevel theory into empirical research. Considerable attention is dedicated to measurement (e.g., measuring environments and individuals within environments) and modeling (e.g., conducting statistical analyses). We emphasize these areas because they are the domains where the least amount of progress has been made and where researchers have called for more novelty and innovation [35, 36].

### Defining and conceptualizing "the environment"

One of the most underappreciated challenges researchers face in translating multilevel research into empirical research relates to how to define and conceptualize the environment. That is, what is the best way to identify environment-level variables or the unique attributes and boundaries of phenomena occurring outside the individual that have an effect upon individuals collectively exposed to such phenomena? What is the appropriate level at which the construct of interest operates? Answers to these questions are critical in determining the types of variability that will be empirically studied.

Traditional definitions of the environment have addressed these questions by emphasizing the physical attributes and resources available to the individual in particular settings and spaces [37]. For example, researchers investigating the effect of neighborhoods on depression and other health outcomes have relied on administrative data, such as census blocks or tracts, counties, or Metropolitan Statistical Areas (MSAs), to define "neighborhood" [38–40]. More recent work has also used "big data" sets, including Google [41]. Similarly, school-based researchers have capitalized on physical delimitations of the space (e.g., school districts, buildings, and classrooms) to define the boundaries of the "school" experience. One of the strengths of these definitions is that they take advantage of general collective agreement about such boundaries; in other words, most people agree on what constitutes School A compared to School B. However, the definition of "neighborhood" can have more subjective meaning [42].

More recent work has defined environment from a transactional perspective, whereby environments are not seen as static entities where groups of individual reside or relate, but rather are defined as a system where transactions between individuals and their social roles are shaped by the available resources and the organizational structure of the setting where those transactions take place [15, 43]. As such, environments are defined in terms of the reciprocal social interplay that exists in the social and temporal space among individuals. The strengths of these transaction-oriented analytic models are that they enable researchers to adopt a less static conceptualization of the environment and recognize the dual role of individuals as both the subject of environmental effects and as active agents in the modification of such environments. However, this transactional approach is methodologically challenging to implement; it requires consideration of how best to operationalize and measure the interplay between individuals and multiple features of their setting.

### Designing the study

The ability of any study to identify causal effects of relationships between children and their environment rests upon its design. In multilevel research, like all other areas of science, the ideal study design to understand causality remains the randomized control trial. Randomization to an environmental condition balances known and unknown confounders across intervention and control groups and prevents selection bias at the level of the environment [44]. Group-level experimental designs finding favorable effects on children's mental health now exist for neighborhood- [45, 46] and school-level studies [47]. Although expensive and challenging to implement, more experimental designs are needed. These designs overcome the challenges that arise in other study designs, which cannot be overcome by even the best analytic techniques.

## Operationalizing and measuring environments and individuals within the environment

### *What is measured?*

In studies investigating multilevel phenomena, deciding what to study is not only an issue of deciding what instrument best measures the construct of interest. Researchers interested in conducting multilevel research must also decide, based on substantive and analytic concerns, at what level the investigation should be conducted. It is thus not only just a matter of picking the right instrument, but it is also a matter of picking the right source of information, at what level, and making thoughtful decisions about what to do with the data so that it represents the environmental characteristics of interest.

### *Who provides data?*

After deciding what will be measured, researchers must then decide *who* is best positioned to report on this characteristic. There is oftentimes a range of potential reporters. For example, in the school setting, potential reporters of the environment-level phenomenon (e.g., school climate) may include students (who can report on their own or their classmates behavior and experiences in school), teachers (who can report on each child's behavior or the behavior of the classroom or school as a unit), or external reporters (who can observe behaviors within the school to rank its school climate). Each reporter has strengths and weaknesses and potentially provides a unique perspective about the phenomena. For example, student reporters may be appropriate to use if students' perceptions of the school climate are considered more salient to the outcome of interest than the perceptions of reporters more distal to the outcome. Students are also appropriate sources of information when the interest is on understanding variation between students, such as in their perceptions or other subjective phenomenon. However, students may be limited in their experience with the broader school environment. Teachers, on the other hand, may be better reporters of school-level practices that are beyond the experience of most students. However, teachers and students may not be the best reporters given that their membership in the school may bias their reports of the school environment. This bias may operate the same or differently across schools and may be of concern when multiple classrooms or schools are studied and the researcher is interested in making comparisons of characteristics across classrooms and/or schools.

One way researchers have addressed this concern has been to use one set of reporters for the predictor variables and another set of reporters for the outcome. For example, neighborhood researchers have used surveys of community residents to ascertain information about the social processes of the neighborhood, including collective efficacy, and then examine associations between collective efficacy and child-level outcomes ascertained from a separate survey [39]. Reporters external to the setting can be used to assess the environment, using objective protocols such as the Classroom Assessment Scoring System [48], designed for observing school classrooms, or systematic social observations of neighborhoods [49]. However, as with teachers and students, external reporters inevitably employ their own frame of reference in observing environments, even those to which they do not belong. Thus, observed variability between environments could be based on biased reporter ratings. Deciding who the best reporter is to describe the environment should therefore be guided by the hypothesized link between what is measured and the outcome. Triangulation methods, or using multiple reporters to describe the same construct of interest, may also be useful, though challenges will arise in deciding how to address discrepancies among multiple reports [50].

Instead of using reporters, investigators have drawn from publicly available data (e.g. census-based, educational authority data). These types of measures are widely used given that they are easily identifiable and can be easily linked to a child's environment. However, publicly available measures often provide good information regarding the objective features of an environment, rather than the subjective domains or social processes. Spatial measures, which define environments as a system involving transactions between individuals and the organizational structure and resources of the setting where those transactions take place, represent one major advancement in measuring environment using publicly available data [15, 43]. However, these approaches are methodologically challenging to implement; they require consideration of how best to operationalize and measure the interplay between individuals and multiple features of their setting. Beyond these measures just described, there has been very little progress in the measuring social environments that may be relevant for understanding the social determinants of mental health.

## Constructing variables

After making decisions regarding what will be measured, researchers must decide on modeling, or how the variables of interest will be organized for use in statistical analyses. Traditionally, data gathered from multilevel analyses have yielded two types of variables. The first, called integral or global variables, refers to inherently environmental-level phenomenon [51]. Integral variables are not derived from the characteristics of individuals within the group and therefore have no individual-level analog. In the context of schools, integral variables may include variables derived from school administrative databases, including school rules and policies (e.g., rules related to attendance or student conduct), school-wide demographic characteristics (e.g., school size, student-to-teacher ratio), teacher characteristics (e.g., teacher quality, percentage of teachers with a master's degree), and school-wide resources (e.g., technology, health services, physical activity facilities). Integral variables may also include variables derived from external rater observations. The primary strength of these variables is that they represent variation in environments measured at the level of the environment itself. Reporters, who are often external to the environment being measured, may be less likely to have biases that would also be linked systematically to the outcome of interest. However, integral variables can also suffer from a lack of direct connection to the individual in the setting. This is of note, as theorists, including Bronfenbrenner [13], argue that an individual's perception of the environment is more important than other qualities.

The second type of variable, called derived variables, is created by summarizing the characteristics of individuals within a group, using means, medians, proportions, measures of dispersion (e.g., variances) or other aggregation approaches [51]. Data to construct derived variables typically come from individual surveys or objective data (e.g., Census-based measures). These variables are also known as composition models [52] or contextual or analytic variables [53]. In the case of schools, derived variables may include mean or median levels of alcohol and drug use, the standard deviation of the distribution of parent income within the school, or the number of students who report depressive symptoms above a predetermined threshold on a survey of such symptoms. Since derived variables are constructed from individual-level data, they are often referred to as aggregates, particularly when measures of central tendency are used.

Like integral variables, derived variables also have their strengths and weaknesses. Derived variables are easy to construct and have been commonly used. However, they do not represent a full translation of multilevel theory for several reasons. First, given how they are calculated (e.g., aggregating individual response to a survey item), they make the assumption that individuals influence their environment, but do not consider how the environment influences individuals. Second, the quality of a derived variable is contingent upon the number of individuals who contribute data. Specifically, the number of individual responses that are being averaged within an environment to create an environment-level mean will likely vary from setting to setting. Thus, environments with a larger number of respondents will have more precise estimates of the environmental-level phenomenon than environments with a smaller number of respondents. Third, different methods of creating derived variables may also lead to different conclusions regarding the effect of the environment on a given outcome. Thus, there is a need for an alternative set of variables that can be used, when either integral or derived variables are not ideal, to measure and model children within their embedded contexts. One alternative approach is described in the third section of the paper.

## Conducting statistical analysis

One obstacle to effectively translating multilevel theory into empirical research is centered upon the lack of available and accessible analytic models that adequately reflect the complexity outlined by these theories, especially with respect to modeling relationships between and among variables at different levels. One analytic strategy that has grown in popularity within the last decade is multilevel modeling (MLM), particularly multilevel generalized linear modeling. As described in detail by others [51, 54–56], MLM provides a broad framework that enables researchers to answer questions about the relationship between children and their environment by explicitly modeling nested relationships. MLM allows for examination of between-child and between-environment differences (or variation in measured characteristics) by incorporating variance parameters. It also allows investigation of differences in the effect of environments on children (or variation in association effects); this can be accomplished by adding an additional parameter (e.g., cross-level interaction term) to the statistical model. MLM can also take into account time, through modeling time as a level.

Although MLM methods are popular, they have limitations, both generally and specifically with respect to translating multilevel theory into empirical research. These limitations are rarely discussed in the literature, which has resulted in an oversimplified view of how to best translate multilevel theories into analytic methods (i.e., the assumption is that multilevel theory equals multilevel model). One of the limitations of MLM is that although it encourages researchers to think about variables as existing at multiple-levels, it typically only does so in relation to the outcome; predictor variables are often assumed to exist at one level. This is apparent in the partitioning of variation in the outcome, but not the predictors, into between- and within-level components. The restriction of only explicitly partitioning the outcome variance forces analysts to create artificial decompositions of variability in select predictors, such as using a derived variable at the between-level in conjunction with the corresponding individual values at the within-level (i.e., when studying neighborhood socioeconomic deprivation, researchers often include a variable capturing neighborhood-level socioeconomic deprivation, along with individual-level socioeconomic status). Moreover, like all regression models, MLM does not allow for a nuanced exploration of the associations between the predictor variables and multiple outcomes; it traditionally focuses on relationships between a set of exogenous predictors and a single outcome, such that the effect of each predictor on the outcome is adjusted for the effects of the other predictors. Thus, it is impossible in traditional MLM to explore bidirectional associations between individuals and environments or to examine the ways in which relationships among the predictor variables can differ across individuals and contexts, for example, examining contextual variability in mediated pathways of influence (i.e., constructs that are presumed to be on the pathway between a predictor and an outcome). Methods are needed that allow for the two core elements of multilevel theory to be more thoroughly empirically studied.

## Opportunities for translating multilevel theory into multilevel research

Within the last decade, novel applications of existing analytic techniques have emerged that provide a flexible analytic alternative to the traditional MLM approaches used by applied researchers. In this section, we describe two of these methods: (1) multilevel factor analysis (MLFA); multilevel structural equation modeling (ML-SEM); and (2) dynamic systems approaches. For each, we summarize the model assumptions, data requirements, and the strengths and limitations with respect to translating multilevel theory. We also provide examples of empirical studies on the social determinants of child mental health that have effectively used the method. Readers interested in other techniques, including cross-classified multilevel models [57] and variations on the longitudinal model [58], are referred elsewhere.

### Multilevel factor analysis

Multilevel factor analysis (MLFA) enables researchers to generate variables that differ from the commonly used integral and derived variables, ideally allowing for better representations of variability in a measured individual or environmental-level characteristic. MLFA is similar to all factor analytic methods in that it seeks to capture the shared variance among an observed or measured set of variables in terms of a potentially smaller number of unobserved constructs or latent factors [59, 60]. Unlike a single-level exploratory or confirmatory factor analysis, which estimates latent factors at only one level (i.e., the child or environmental level), MLFA decomposes the total sample variance–covariance matrix into child-level (within an environment) and environment-level matrices and simultaneously models *distinct* latent factor structures at each of these levels [61–63]. Thus, the result is a measurement model that relates each item to a latent factor, at both the individual and environment levels. This measurement model consists of three parameters: (1) intercept; (2) factor loadings (e.g., the slope or coefficient relating the underlying latent factor to the observed variable); (3) residual variance (e.g., the unique variance in the observed indicator variables not explained by or related to the latent factor; this unique variance is a combination of measurement error and other sources of variability).

Conventional MLFA operates under the same assumptions as linear regression, including multivariate normality and Homoscedasticity; however, just as in logistic regression, these assumptions are not required to be met when categorical data are analyzed [59]. The tenability of these assumptions can be evaluated using careful data screening. These assumptions can also be relaxed using alternative link functions and error distributions (e.g., generalized linear models).

MLFA allows for both theory-testing and theory-generation. Detailed a priori knowledge of the factor structure underlying the item set is not necessary, as a multilevel exploratory factor analysis (ML-EFA) will provide information about the number of latent factors underlying an item set and the relationship between factors and observed variables. Thus, a ML-EFA can be helpful when there is a lack of detailed theory regarding the constructs of interest. Following ML-EFA, a multilevel confirmatory factor analysis (ML-CFA) should be conducted. In ML-CFA, the number of factors and relationship between factors and indicators are known and the goal is to validate this hypothesized model. Thus, ML-CFA can be useful for theory-testing. MLFA can accommodate sampling weights [64] and multiple data types (e.g., ordinal, categorical), and can be used when multiple data sources are used (e.g., multiple reporters, administrative data) thus allowing for triangulation of measures.

To date, most work on MLFA has been published in methodological, rather than applied, journals. However, a growing number of studies are emerging, particularly in education and psychology [see for example 65, 66] that provide an application of MLFA methods. For example, a study by Dunn and colleagues provides the first practical illustration of the MLFA method for use by public health-oriented multilevel researchers [67]. In this study, the authors used MLFA to examine the construct collective efficacy. Collective efficacy was first introduced by Sampson and colleagues as a feature of neighborhoods that consists of two dimensions: social cohesion among neighbors (social cohesion) and neighbors' willingness to intervene on behalf of the common good (informal social control) [68]. Dunn and colleagues used MLFA to examine collective efficacy because despite being one of the most popular constructs studied in social epidemiological research, few papers (outside of the original paper by Sampson) have used latent variable approaches to study collective efficacy. What Dunn and colleagues found was that the best fitting MLFA model was one that modeled collective efficacy as a two-dimensional construct at the within-level, consisting of the two latent constructs informal social control and social cohesion, and a one-dimensional construct at the between-level, consisting of collective efficacy. Thus, the finding that there were different latent factor structures at each level (individual and neighborhood) led them to conclude that there is a need to separately consider and measure phenomenon at each level of analysis. Although interesting, these results must be replicated in future studies.

There are several advantages of MLFA, both generally and specifically in relation to translating multilevel theories into multilevel research. First, MLFA allows for variation in latent factors at the child and environment levels, which is one of the core elements emphasized by multilevel theories. By modeling two different latent factor structures (at the child and environment level), MLFA allows researchers to better understand the variation in structure and meaning that is driven by child-level constructs that vary across children within an environment and environment-level constructs that are the same across children (in a shared environment) but that vary between environments. Thus, MLFA is distinct from a model that merely decomposes the variability of individual-level factors into child- and environment-level components or presumes the same underlying factor structure across levels. Second, using a MLFA specification for predictors as well as outcome variables enables the variability in predictors of interest to be separated into child- and environment-level components during the model estimation, eliminating the need to use derived variables. Third, the MLFA method (particularly ML-EFA) can lead to the development of new theories, as it gives researchers the opportunity to use existing data to identify potentially new constructs, especially at the environment level and in relation to environments that have not yet been widely studied, including schools. One of the major types of new constructs that could be identified through these methods is emergent properties, or the characteristics of the environment that arise from exchanges between people in a given setting. Thus, MLFA may allow for a better translation of multilevel theory into multilevel research through the identification of new and potentially more relevant environment-level phenomenon and subsequently to the identification of new measures of the social environment.

Although a potentially promising method, MLFA is certainly not without its limitations. First, the MLFA method can be computationally intense, depending on the sample size and number of items used. Thus, to effectively use MLFA methods, analysts may need to limit the number of items included in the analysis, thus inhibiting the potential of the method. Analysts can analyze data using high-performance computing facilities, though access to these kinds of resources may be limited. Moreover, MLFA methods are best performed using specialized software (e.g., MPlus, LISREL), which may be inaccessible to some applied researchers with limited funding resources. Some of these software packages only allow for two-level structures (e.g., adolescents nested in neighborhoods), making examination of three-level structures (e.g., repeated measures nested in adolescents, nested in neighborhoods) impossible. Finally, factor analytic techniques do require researchers to balance both empirical fit with theoretical understanding, which can lead to the generation of different results (and interpretation of them). Additional studies on MLFA will be needed to evaluate the extent to which this method provides multilevel researchers with a unique tool to study environments and identify characteristics of the social environment that may be relevant for children's mental health.

### Multilevel structural equation modeling

Multilevel structural equation modeling (ML-SEM) is an outgrowth of single-level structural equation modeling (SEM) [59, 69, 70]. In a single-level SEM, one or more measurement models (i.e., results from a MLFA) are joined together in a structural model, where associations between latent variables, covariates, and observed variables are estimated. These associations are estimated using a path model, which estimates direct (non-mediated) and indirect (mediated) paths between variables or latent factors; parameters from the path model are identical to a multiple regression. In ML-SEM, a structural model exists at two or more levels (i.e., child and environment), thus allowing for estimation of associations among variables and latent factors at multiple levels of analysis (e.g., how neighborhood latent factors predict neighborhood levels of child depressive symptoms). As in all MLM analyses, ML-SEM can include random intercepts (e.g., differences in the average outcome for each neighborhood), random slopes (e.g., differences in the effect of neighborhood characteristics on child depressive symptoms by neighborhood), and cross-level interactions (e.g., differences in the average effect of a neighborhood characteristic on child depressive symptoms for girls compared to boys). Thus, ML-SEM allows for investigation of variation in both measured characteristics and variation in the effect of environments on individuals, as suggested by multilevel theorists.

Assumptions for ML-SEM are the same as any regression. Many software packages used to carry out ML-SEM can also accommodate imputed data to address missingness [71]. Although data for ML-SEM can be cross-sectional, it is ideally suited to prospectively collected, longitudinal data, given its ability to model complex association pathways linking multiple predictor variables to other predictor variables and then outcomes. Indeed, longitudinal data with multiple measurement points are the only way to truly test the reciprocal determinism concept.

The number of studies using ML-SEM are growing. For example, recent applications of ML-SEM include a longitudinal ML-SEM model examining children at risk for eating disorders [72], a study on neighborhood effects on adolescent's sense of self-control [73], and the effect of alcohol outlet density on adolescent's alcohol use [74]. However, like MLFA, most work on ML-SEM has been published in statistical methods journals. Applications of ML-SEM have also been limited largely to organizational psychology, samples from a select set of countries, and have not included a measurement model with different latent factor structures at the child and environment level.

ML-SEM has several advantages. With respect to translating multilevel theory and its core elements into empirical research, one of the real strengths of ML-SEM is that it allows for the examination of direct and indirect pathways of influence within and across levels of influence. In other words, ML-SEM enables researchers to incorporate, into one analysis, latent variables that can simultaneously be both outcomes and predictors. This is accomplished because ML-SEM allows for feedback loops, specifically in the form of multivariate outcomes; it models relationships *among* the predictor variables, and allows both direct and indirect pathways from predictors to the different outcomes to vary across environments. As a result, ML-SEM provides a flexible analytic tool that enables researchers to examine pathways linking environmental conditions to individual mental health and thus test for multilevel mediation [75]. The ideal scenario for such tests of mediation is in the context of longitudinal data, when the exposure is temporally prior to the mediator and both are temporally prior to the outcome. ML-SEM can also allow for investigation of multidirectional associations or multiple pathways linking exposure to outcome [76] and the reciprocal relationships between variables emphasized by multilevel theories. Thus, ML-SEM provides the tools to empirically test complex etiological models.

Finally, one of the main general advantages of ML-SEM is that it can partially account for some measurement error in the predictor. Unlike regression models, which presume that there is no error in the measurement of the predictor, the measurement model provided by the MLFA includes a parameter that explicitly captures unique variance, which as noted previously includes both random (e.g., measurement) and non-random error. Thus, while not completely removed, bias in the predictor variable will be partially removed through the error term in ML-SEM as described by DeShon [77]. However, one of the major disadvantages of ML-SEM is that without longitudinal data, it is impossible to model the reciprocity between children and their environments. Moreover, even when longitudinal data are available, the temporal ordering of constructs must be carefully measured and specified to allow for the accurate assessment of the temporal processes under study.

## Dynamic systems

System modeling methodologies, including agent-based modeling, capitalize on prior empirical work using large-scale databases to construct prediction models linking individual- and environment-level influences on mental health and other outcomes [78]. Using a set of mathematical equations, these simulation models describe the relationships between different variables that directly or indirectly influence an outcome of interest. These models explicitly characterize relationships between different social actors in an environment, thus allowing for examination of the interrelationships and reciprocal nature of mental health and social outcomes not easily studied within a traditional statistical model. A growing number of studies are using these techniques, particularly in public health (see for example the 2006 special issue of the American Journal of Public Health). Additionally, recent work has applied these methods to the study of child development [79, 80], obesity [81], and drug use [82].

One of the major strengths of dynamic systems models is that they allow researchers to experimentally manipulate conditions that could be the target of new policies or interventions, but which may not be easily or ethically adaptable in human populations. For example, a dynamic systems model could study and manipulate constructs such as the supply of medication or the prevalence of particular risk behaviors in a given population. Dynamic systems models are therefore useful in providing a virtual platform for testing competing hypotheses regarding the importance of different kinds of variability; this virtual platform ensures no risk to human subjects, no waste of physical resources, and an opportunity to repeat experiments in a short span of time. However, the appropriate specification of these models requires in-depth content knowledge so that all relevant individual and environmental determinants are included; this expertise is often best obtained through an interdisciplinary team [83]. A high level of technological expertise is also required; improper specification of these models is a huge concern as it may lead intelligent but naïve users to make incorrect inferences from the model [84]. When used carefully, dynamic systems models may reveal new pathways for intervention and determine new directions for empirical work with human subjects.

## Conclusion

Deeply shaped by the context in which they are articulated, theories offer researchers an organizing framework from which to generate research questions, develop testable hypotheses, and ultimately understand and reduce urgent health issues [85, 86]. Theories have both propelled and impeded understanding of health problems and the development of appropriate policies and interventions to reduce them; this is evident today as much as it was centuries ago [87] and is apparent in the issue of how best to translate multilevel theories into empirical research.

This paper sought to raise awareness regarding the challenges that exist in bridging multilevel theory and multilevel research, particularly as they relate to understanding the social determinants of child mental health. By making this translational gap explicit, the field can make progress towards building bridges between theory and research. Many barriers prevent researchers from effectively bridging the multilevel theory to multilevel research translational gap. These barriers exist at all stages of the research process, though the areas of measurement and analysis represent the areas with the greatest potential for innovation.

A major take-home of this paper is that new methodological tools and techniques are needed to answer complex questions related to child mental health and better understand children in context. The field requires better translation of existing methods, including generalized multilevel modeling, as well as emerging methods, including those described here (e.g., MLFA and ML-SEM). Researchers need resources that provide accessible and clear direction on how to apply and interpret the findings of methodologically sophisticated techniques. Finally, scientific knowledge must continue to evolve circuitously, with methodological advancements informing and refining theory, and new developments in theory informing the development of new methods. Thus, there is a need for both theory development and refinement of theory based on theory testing.

To achieve real advancements in promoting child mental health and understanding the social determinants of child mental health, our field must grapple with theoretical and methodological complexity. It is time that our current technological capabilities join forces with our theories to reduce the multilevel theory to empirical research translational gap. By building more complete bridges between multilevel theory and multilevel research, the field can bring about a new generation of studies that use innovative empirical methods to realize the full potential of multilevel theories and lead to new discoveries about the social determinants of child mental health.

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

## References


1. Sigerist HE (1961) A history of medicine, vol II. Early greek, hindu, and persian medicine. Oxford University Press, New York
2. Villerme LR (1988) A description of the physical and moral state of workers employed in cotton, wool, and silk mills. In: Buck C et al (eds) The challenge of epidemiology: issues and selected readings. PAHO/WHO, Washington, DC, pp 33–36
3. Engels F (1958) The condition of the working class in england. Stanford University Press, Stanford (Transl: OW Henderson, WH Chalones, 1848)
4. Virchow L (1988) Report on the typhus epidemic in upper silesia. In: Rather LJ (ed) Rudolf Virchow: collected essays on public health and epidemiology. Science History Publications, Canton, pp 205–319
5. Cassel J (1976) The contribution of the social environment to host resistance. Am J Epidemiol 104:107–123
6. Doyal L (1979) The political economy of health. Pluto Press, London
7. Krieger N (2001) Theories for social epidemiology in the 21st century: an ecosocial perspective. Int J Epidemiol 30:668–677
8. Stokols D (1996) Translating social ecological theory into guidelines for community health promotion. Am J Health Promot 10:282–298
9. Dahlgren G, Whitehead M (1991) Policies and strategies to promote social equity in health: Background document to WHO—Strategy paper for Europe. Institute for Futures Studies, Stockholm
10. Bandura A (1986) Social foundations of thought and action: a social cognitive theory. Prentice-Hall Inc, Englewood Cliffs
11. Sameroff A, Chandler R (1975) Reproductive risk and the continuum of caretaking casualty. In: Horowitz FD (ed) Editor review of child development research. University of Chicago Press, Chicago, pp 187–244

12. Sameroff A (2010) A unified theory of development: dialectic integration of nature and nurture. Child Dev 81(1):6–22
13. Bronfenbrenner U (1979) The ecology of human development. Harvard University Press, Cambridge
14. Bronfenbrenner U, Morris PA (2006) The bioecological model of human development. In: Lerner RM (ed) Handbook of child psychology. Wiley, Hoboken, pp 793–828
15. Tseng V, Seidman E (2007) A systems framework for understanding social settings. Am J Community Psychol 28:217–228
16. Goodman E et al (2007) Perceived socioeconomic status: a new type of identity that influences adolescents' self-rated health. J Adolesc Health 41:479–487
17. Caspi A et al (2010) Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. Am J Psychiatry 167(5):1–19
18. Dunn EC et al (2011) Gene–environment interaction (GxE) research in youth depression: a systematic review with recommendations for future research. J Child Psychol Psychiatry 52(12):1223–1238
19. Boyce WT, Ellis BJ (2005) Biological sensitivity to context: i. An evolutionary-developmental theory of the origins and functions of stress reactivity. Dev Psychopathol 17:271–301
20. McLaughlin KA et al (2010) Developmental origins of emotional reactivity to stress and life course associations with mood and anxiety disorders. Depression Anxiety 27:1087–1094
21. Luthar SS, Sawyer JA, Brown PJ (2006) Conceptual issues in studies of resilience: past, present, and future research. Ann N Y Acad Sci 1094:105–115
22. Plomin R, Daniels D (2011) Why are children in the same family so different from one another? Int J Epidemiol 40:563–582
23. Kendler KS, Baker JH (2007) Genetic influences on measures of the environment: a systematic review. Psychol Med 37:615–626
24. Rutter M et al (2001) Testing hypotheses on specific environmental causal effects on behavior. Psychol Bull 127:291–324
25. Scarr S, McCartney K (1983) How people make their own environments: a theory of genotype-environment effects. Child Dev 54:424–435
26. Masten AS, Cicchetti D (2010) Developmental cascades. Dev Psychopathol 22:491–495
27. Lloyd JEV, Li L, Hertzman C (2010) Early experiences matter: lasting effect of concentrated disadvantage on children's language and cognitive outcomes. Health Place 16:371–380
28. Dunn EC, McLaughlin KA, Slopen N, Rosand J, Smoller JW (2013) Developmental timing of child maltreatment and symptoms of depression and suicidal ideation in young adulthood: results from the National Longitudinal Study of Adolescent Health. Depress Anxiety 30:955–964
29. Berkman LF, Kawachi I (2000) Social epidemiology. Oxford University Press, New York
30. Leach LS et al (2013) Relationship quality and levels of depression and anxiety in a large population-based survey. Soc Psychiatry Psychiatr Epidemiol 48(3):417–425
31. Belsky J, Pleuss M (2009) Beyond diathesis stress: differential susceptibility to environmental influences. Psychol Bull 135:885–908
32. Ellis BJ, Boyce WT (2008) Biological sensitivity to context. Curr Dir Psychol Sci 17:183–187
33. Richardson J et al (2012) The influence of state mental health perceptions and spending on an individual's use of mental health services. Social Psychiatry Psychiatric Epidemiol
34. Sampson RJ, Raudenbush SW (2004) Seeing disorder: neighborhood stigma and the social construction of "Broken Windows". Social Psychol Q 67:319–342
35. Mujahid MS et al (2007) Assessing the measurement properties of neighborhood scales: from psychometrics to ecometrics. Am J Epidemiol 165(8):858–867
36. Raudenbush SW, Sampson RJ (1999) Ecometrics: toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. Sociol Methodol 29:1–41
37. Duncan C, Jones K (1993) Do places matter? A multi-level analysis of regional variation in health related behavior in Britain. Soc Sci Med 37:725–733
38. Mair CF, Diez Roux AV, Galea S (2008) Are neighborhood characteristics associated with depressive symptoms? A critical review. J Epidemiol Community Health 62(11):940–946
39. Xue Y et al (2005) Neighborhood residence and mental health problems of 5- to 11-year olds. Arch Gen Psychiatry 62:554–563
40. Kim D (2008) Blues from the neighborhood? Neighborhood characteristics and depression. Epidemiol Rev 30:101–117
41. Odgers CL et al (2012) Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. J Child Psychol Psychiatry 53(10):1009–1017
42. Campbell E et al (2009) Subjective constructions of neighborhood boundaries: lessons from a qualitative study of four neighborhoods. J Urban Aff 31(4):461–490
43. Cummings S et al (2007) Understanding and representing "place" in health research: a relational approach. Soc Sci Med 65:1825–1838
44. Murray DM (1998) Design and analysis of group-randomized trials. Oxford University Press, New York
45. Leventhal T, Brooks-Gunn J (2003) Moving to opportunity: an experimental study of neighborhood effects on mental health. Am J Public Health 93(9):1576–1582
46. Kling JR, Liebman JB, Katz LF (2007) Experimental analysis of neighborhood effects. Econometrica 75(1):83–119
47. Jones SM, Brown JL, Aber JL (2011) Two-year impacts of a universal school-based social-emotional and literacy intervention: an experiment in translational developmental research. Child Dev 82(2):533–554
48. Pianta RC, La Paro K, Hamre BK (2008) Classroom assessment scoring system (CLASS). Paul H. Brookes, Baltimore
49. Sampson RJ, Raudenbush SW (1999) Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. Am J Sociol 105(3):603–651
50. Kraemer HC et al (2003) A new approach to integrating data from multiple informants in psychiatric assessment and research: mixing and matching contexts and perspectives. Am J Psychiatry 160(9):1566–1577
51. Diez Roux AV (2002) A glossary for multilevel analysis. J Epidemiol Community Health 56:588–594
52. Chan D (1998) Functional relations among constructs in the same content domain at different levels of analysis: a typology of composition models. J Appl Psychol 83(2):234–246
53. Ludtke O et al (2008) The multilevel latent covariate model: a new, more reliable approach to group-level effects in contextual studies. Psychol Methods 13(3):203–229
54. Diez Roux AV (1998) Bringing context back into epidemiology: variables and fallacies in multilevel analysis. Am J Public Health 88(2):216–222
55. Raudenbush SW, Bryk AS (2002) Hierarchical linear models: applications and data analysis method, 2nd edn. Sage Publications, Thousand Oaks
56. Subramanian SV, Jones K, Duncan C (2003) Multilevel methods for public health research. In: Kawachi I, Berkman LF (eds) Neighborhoods and health. Oxford University Press, New York, pp 65–111
57. Goldstein H (1994) Multilevel cross-classified models. Sociol Methods Res 22:364–375
58. Singer JD, Willett JB (2003) Applied longitudinal data analysis: modeling change and event occurrence. Oxford University Press, New York

59. Kline RB (2010) Principles and practice of structural equation modeling, 3rd edn. Guilford Press, New York
60. Brown TA (2006) Confirmatory factor analysis for applied research. Guilford Press, New York
61. Muthen BO (1991) Multilevel factor analysis of class and student achievement components. J Educ Meas 28(4):338–354
62. Muthen BO (1994) Multilevel covariance structure analysis. Sociol Methods Res 22:376–398
63. Hox JJ (2010) Multilevel analysis: techniques and applications, 2nd edn. Routledge, New York
64. Muthén LK, Muthén BO (2007) Mplus user's guide, 6th edn. Muthén & Muthén, Los Angeles
65. Dedrick RF, Greenbaum PE (2011) Multilevel confirmatory factor analysis of a scale measuring interagency collaboration of children's mental health agencies. J Emot Behav Disord 19:27–40
66. Reise SP et al (2005) An illustration of multilevel factor analysis. J Pers Assess 84(2):126–136
67. Dunn EC et al Modeling contextual effects using individual-level data and without aggregation: an illustration of multilevel factor analysis (MLFA) with collective efficacy (unpublished)
68. Sampson RJ, Raudenbush S, Earls F (1997) Neighborhoods and violent crime: a multilevel study of collective efficacy. Science 277:918–924
69. Marsh HW et al (2009) Doubly-latent models of school contextual effects: integrating multilevel and structural equation approaches to control measurement and sampling error. Multivar Behav Res 44:764–802
70. MacCallum RC, Austin JT (2000) Applications of structural equation modeling in psychological research. Annu Rev Psychol 51:201–226
71. Tomarken AJ, Waller NG (2005) Structural equation modeling: strengths, limitations, and misconceptions. Annu Rev Clin Psychol 1:31–65
72. Allen KL, Byrne SM, McLean NJ (2012) The dual-pathway and cognitive-behavioural model of binge eating: prospective evaluation and comparison. Eur J Child Adolesc Psychiatry 21:51–62
73. Teasdale B, Silver E (2009) Neighborhoods and self-control: toward an expanded view of socialization. Soc Probl 56(1):205–222
74. Kuntsche E, Kuendig H, Gmel G (2008) Alcohol outlet density, perceived availability and adolescent alcohol use: a multilevel structural equation model. J Epidemiol Community Health 62(9):811–816
75. Preacher KJ, Zyphur MJ, Zhang Z (2010) A general multilevel SEM framework for assessing multilevel mediation. Psychol Methods 15(3):209–233
76. Galea S, Ahern J (2006) Invited commentary: considerations about specificity of associations, causal pathways, and heterogeneity in multilevel thinking. Am J Epidemiol 163(2):1079–1082
77. DeShon RP (1998) A cautionary note on measurement error corrections in structural equation models. Psychol Methods 3(4):412–423
78. Homer JB, Hirsch GB (2006) System dynamics modeling for public health: background and opportunities. Am J Public Health 96(3):452–458
79. Urban JB, Osgood ND, Mabry PL (2011) Developmental systems science: exploring the application of systems science methods to developmental science questions. Res Human Dev 8(1):1–25
80. Cox MJ et al (2010) Systems theory and cascades in developmental psychopathology. Dev Psychopathol 22(3):497–506
81. Galea S, Riddle M, Kaplan GA (2010) Causal thinking and complex system approaches in epidemiology. Int J Epidemiol 39:97–106
82. Galea S, Hall C, Kaplan GA (2009) Social epidemiology and complex system dynamic modelling as applied to health behaviour and drug use research. Int J Drug Policy 20:209–216
83. Mabry PL et al (2008) Interdisciplinarity and systems science to improve population health: a view from the NIH Office of Behavioral and Social Sciences Research. Am J Prev Med 35(2S):S211–S224
84. Sterman JD (2006) Learning from evidence in a complex world. Am J Public Health 96(3):505–514
85. Levins R (1996) Ten propositions on science and antiscience. Social Text 46–47(14):101–111
86. Ziman J (2000) Real science: what it is and what it means. Cambridge University Press, Cambridge, pp 126–134, 144–151
87. Fleck L (1979) Genesis and development of a scientific fact. University of Chicago Press, Chicago