Journal of Youth and Adolescence
https://doi.org/10.1007/s10964-023-01767-w

EMPIRICAL RESEARCH

Predicting Adolescent Mental Health Outcomes Across Cultures: A Machine Learning Approach

W. Andrew Rothenberg[1,2] · Andrea Bizzego[3] · Gianluca Esposito[3] · Jennifer E. Lansford[1] · Suha M. Al-Hassan[4] · Dario Bacchini[5] · Marc H. Bornstein[6,7] · Lei Chang[8] · Kirby Deater-Deckard[9] · Laura Di Giunta[10] · Kenneth A. Dodge[1] · Sevtap Gurdal[11] · Qin Liu[12] · Qian Long[13] · Paul Oburu[14] · Concetta Pastorelli[10] · Ann T. Skinner[1] · Emma Sorbring[11] · Sombat Tapanya[15] · Laurence Steinberg[16,17] · Liliana Maria Uribe Tirado[18] · Saengduean Yotanyamaneewong[15] · Liane Peña Alampay[19]

Received: 13 October 2022 / Accepted: 13 March 2023

Abstract
Adolescent mental health problems are rising rapidly around the world. To combat this rise, clinicians and policymakers need to know which risk factors matter most in predicting poor adolescent mental health. Theory-driven research has identified numerous risk factors that predict adolescent mental health problems but has difficulty distilling and replicating these findings. Data-driven machine learning methods can distill risk factors and replicate findings but have difficulty interpreting findings because these methods are atheoretical. This study demonstrates how data- and theory-driven methods can be integrated to identify the most important preadolescent risk factors in predicting adolescent mental health. Machine learning models examined which of 79 variables assessed at age 10 were the most important predictors of adolescent mental health at ages 13 and 17. These models were examined in a sample of 1176 families with adolescents from nine nations. Machine learning models accurately classified 78% of adolescents who were above-median in age 13 internalizing behavior, 77.3% who were above-median in age 13 externalizing behavior, 73.2% who were above-median in age 17 externalizing behavior, and 60.6% who were above-median in age 17 internalizing behavior. Age 10 measures of youth externalizing and internalizing behavior were the most important predictors of age 13 and 17 externalizing/internalizing behavior, followed by family context variables, parenting behaviors, individual child characteristics, and finally neighborhood and cultural variables. The combination of theoretical and machine-learning models strengthens both approaches and accurately predicts which adolescents demonstrate above average mental health difficulties in approximately 7 of 10 adolescents 3–7 years after the data used in machine learning models were collected.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s10964-023-01767-w.

W. Andrew Rothenberg
rothenbergdrew@gmail.com

1 Duke University, Durham, NC, USA
2 University of Miami, Coral Gables, FL, USA
3 University of Trento, Trento, Italy
4 Hashemite University, Zarqa, Jordan
5 University of Naples "Federico II", Naples, Italy
6 Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, Maryland, USA
7 UNICEF, New York, New York, USA
8 University of Macau, Zhuhai, China
9 University of Massachusetts, Amherst, MA, USA
10 Università di Roma "La Sapienza", Rome, Italy
11 University West, Trollhättan, Sweden
12 Chongqing Medical University, Chongqing, China
13 Duke Kunshan University, Suzhou, China
14 Maseno University, Maseno, Kenya
15 Chiang Mai University, Chiang Mai, Thailand
16 Temple University, Philadelphia, PA, USA
17 King Abdulaziz University, Jeddah, Saudi Arabia
18 Universidad de San Buenaventura, Bogotá, Colombia
19 Ateneo de Manila University, Quezon, Philippines

Published online: 19 April 2023

Keywords Machine learning · Externalizing · Internalizing · Adolescence · Prediction · Parenting

Introduction

Prevention scientists, clinicians, and policymakers alike have sounded the alarm that the rise in adolescent mental health problems is a global crisis (Benton et al., 2021). Unfortunately, this crisis has only been deepened by the COVID-19 pandemic, as epidemiologists estimate that the prevalence of adolescent internalizing disorders (i.e., depression, anxiety, and other mood difficulties) have doubled from 10% to 20% (Racine et al., 2021) and the prevalence of externalizing problems (i.e., conduct disorders, oppositional defiant disorders, emotional dysregulation and attention disorders) has increased 120%, from 6.7% to 8.1% (Lebrun-Harris et al., 2022). Myriad environmental and contextual risk factors have been identified as contributing to this adolescent mental health crisis, including poor mental health in childhood (Lansford et al., 2018a), numerous maladaptive parenting behaviors (Pinquart, 2017a), adverse family environments (Rothenberg, 2019), dangerous neighborhoods (Leventhal & Dupéré, 2019), and larger cultural norms (Lansford et al., 2018b) that reinforce mental health problems. Though knowledge of these myriad risk factors has deepened understanding of the etiology of adolescent mental health problems (Tate et al., 2020), it also poses a problem. In the midst of this "information overload," how can interventionists and policymakers identify which risk factors matter most in predicting poor adolescent mental health, and when in adolescence these risk factors matter (Tate et al., 2020)? Distilling the numerous risk factors examined in existing literature to identify what risk factors matter most and when is essential. Doing so allows interventionists and policymakers to identify which interventions targeting specific risk factors might be best to invest in, given finite time and resources (Tate et al., 2020). This study attempts to answer this question by integrating theory-driven and data-driven scientific frameworks to identify which among 79 predictors at the beginning of adolescence are the most powerful predictors of adolescent mental health in mid-adolescence and late-adolescence in a sample of 1176 adolescents from 12 ethnocultural groups in 9 countries.

Two Methods For Identifying the Most Powerful Predictors of Adolescent Mental Health

Prevention scientists have devised two types of methods to answer the question of what risk factors matter the most in predicting adolescent mental health problems: older (and still predominantly used) theory-driven methods (Bronfenbrenner & Morris, 2006) and newer data-driven machine-learning prediction methods (Dwyer et al., 2018). Theory-driven methods generate broad theoretical frameworks that might explain how adolescent mental health problems develop, and then generate empirically testable hypotheses that support or refute those frameworks. For instance, a predominant theoretical framework in contemporary developmental science is the Bioecological Model of Human Development (Bronfenbrenner & Morris, 2006), which posits that youth health (including mental health) develops within a system of multiple developmental levels that constantly interact with one another (Bornstein & Rothenberg, 2022). Specifically, children exhibit their own individual characteristics, which are embedded within and informed by parenting within the parent-child dyad (Lansford et al., 2018a), which is embedded within the family environment (Rothenberg, 2019), which is embedded within the neighborhood context (Leventhal & Dupéré, 2019), which in turn is embedded within a cultural group (Lansford et al., 2018b). Driven by the Bioecological Model, developmental science has made progress in the last 50 years in identifying the child characteristics, parenting and family environments, neighborhood contexts, and cultural norms that influence adolescent mental health (progress that will be reviewed in further depth below; Bronfenbrenner & Morris, 2006).

However, theory-driven methods have run into two obstacles that have been identified in recent years. First, due to limits in computational power and analytic flexibility (Dwyer et al., 2018), most theory-driven analyses have investigated predictors of adolescent mental health at only one or two of these individual, parent, family, neighborhood, and cultural levels of development (Sun et al., 2020). This is part of what has led to the aforementioned "information overload" where many different studies investigating only one or two risk factors at a time have led to a literature rife with numerous predictors of adolescent mental health, but less knowledge of how important those factors are compared to one another (Tate et al., 2020). Second, in recent years the replicability of much theory-driven research in psychology has been called into question (Dwyer et al., 2018). This is because, often times, research is more likely to be published in scientific journals if it finds significant results that support a theory, and less likely to be published in scientific journals if it finds null results that do not support a theory (Dwyer et al., 2018). Therefore, scientists (often unwittingly) engage in methods such as *p*-hacking to ensure that their results are significant and publishable (Dwyer et al., 2018). Consequently, many results in theory-driven research are difficult to replicate, leading to the "replication crisis" in psychology (Dwyer et al., 2018).

To overcome these obstacles, prevention scientists have employed a second set of methods more recently: data-driven machine learning prediction methods (Dwyer et al., 2018). Machine learning algorithms do not have the same limits on computational power and analytic flexibility of more traditional statistical techniques often employed in theory-based work (Sun et al., 2020). Therefore, instead of only investigating one or two predictors of adolescent mental health at a time, data-driven machine learning prediction methods can simultaneously analyze, for instance, 474 predictors of adolescent mental health (Tate et al., 2020) or 298 predictors of adolescent subjective well-being (Zhang et al., 2019) to identify the most powerful among these predictors. This is a promising analytic solution to the information overload problem. Additionally, machine learning algorithms are, by definition, atheoretical (Dwyer et al., 2018). Machine learning algorithms are designed to discern how the predictors provided to them can be combined to create the most accurate predictions of future events (Dwyer et al., 2018). In so doing, machine learning algorithms split datasets into two subsets, a "train" subsample where the machine learning algorithm is trained to optimize prediction, and then a "test" subsample where the trained algorithm is put to the "test" by applying it to new data it has not "seen" before. This data-splitting method empirically evaluates how replicable and generalizable machine learning algorithms are, providing a powerful analytic solution to the replicability crisis.

However, the strengths of machine learning models are also their biggest weakness: their atheoretical nature (Dwyer et al., 2018). Indeed, because machine learning algorithms are not driven by theory, and are instead designed to detect and predict patterns, machine learning results have often been described as "black boxes" that are difficult to interpret or understand (Dwyer et al., 2018). For instance, it is often unknown why certain predictors emerge as especially powerful in machine learning analyses, because no a priori theory has been posited to help interpret the findings (Dwyer et al., 2018).

In sum, theory-driven research (like that of the Bioecological Model) has identified a vast array of risk factors that might predict adolescent mental health but suffers from difficulties distilling those risk factors that are most important and replicating findings (Tate et al., 2020). Data-driven machine learning algorithms are able to distill the vast array of risk factors down to those that are most important in predicting adolescent mental health and do so in a replicable way, but suffer from difficulties with interpretation due to a lack of theoretical support (Dwyer et al., 2018). The present study attempts to integrate these two approaches to accentuate each of their strengths and mitigate their weaknesses.

Using the Bioecological Model to Identify The Most Important Predictors of Adolescent Mental Health

The Bioecological Model of Human Development posits that youth mental health develops within a system of multiple developmental levels that constantly interact with one another (Bornstein & Rothenberg, 2022). At the heart of this model is a youth's individual characteristics (e.g., individual predispositions, abilities, and past health; Bronfenbrenner & Morris, 2006). These individual characteristics are in turn embedded within, and their development is informed by, the *microsystem* that surrounds them (e.g., those aspects of the youth's environment that have direct contact with the child, like their parents and direct family context; Bronfenbrenner & Morris, 2006). This microsystem is, in turn, embedded in, and its development is informed by, the *exosystem* that surrounds it (e.g., those aspects of the larger environment that impact the child's microsystem, like how the safety of a neighborhood impacts how strictly parents parent their children; Leventhal & Dupéré, 2019). This exosystem is, in turn, embedded within a *macrosystem* (e.g., the norms and ideologies of a culture, like individualism and collectivism, which might dictate how safe neighborhoods in given cultural contexts are, or how parents parent in specific cultures; Bornstein et al., 2021). Finally, these systems are embedded in a *chronosystem* (e.g., the passage of time that might lead to environmental changes over ontogeny; Bronfenbrenner & Morris, 2006).

This investigation uses the Bioecological Model's systems framework to organize the 79 age 10 predictors that it uses to prospectively predict adolescent mental health problems at ages 13 and 17. Specifically, this investigation examines two categories of individual characteristics (age 10 child mental health problems and other age 10 individual characteristics such as intelligence, self-regulation, pubertal status, age, gender), two categories of *microsystem* risk factors (parenting behaviors and family context variables), one category of *exosystem* risk factors (family member perceptions of neighborhood safety), and one category of *macrosystem* risk factors (mother and father cultural orientation towards *individualism* or *collectivism*) to determine which of these categories produces the most powerful predictors of adolescent externalizing and internalizing mental health problems. The current investigation incorporates the *chronosystem* by examining whether these age 10 risk factors predict adolescent mental health at age 17 as well as they do at age 13. Myriad evidence from existing theory-driven studies links each of the categories of risk factors with adolescent mental health.

Individual child characteristics

A substantial body of theory-based evidence links the two categories of individual characteristics studied here with adolescent mental health. For instance, several meta-analyses indicate that the strongest predictor of future youth externalizing (Pinquart, 2017a) or internalizing (Pinquart, 2017b) problems is past youth externalizing or internalizing problems. In addition, numerous other youth characteristics, including self-regulation (Strauman, 2017), executive functioning (Bloeman et al., 2018) and pubertal development (Belsky et al., 2020) have also been meta-analytically or longitudinally associated with youth externalizing and internalizing behavior. Greater difficulties with self-regulation and executive functioning and off-track pubertal development are all associated with greater youth internalizing and externalizing behavior (Belsky et al., 2020).

Microsystem risk factors

Theory-based work has also linked both categories of microsystem factors examined in this study (parenting variables and family context variables) with the development of youth mental health problems. With regards to parenting, two recent meta-analyses have demonstrated that a suite of parenting behaviors including low warmth, low behavioral control, low autonomy granting, high harsh parenting, and high psychological control predicted greater adolescent externalizing (Pinquart, 2017a) and internalizing (Pinquart, 2017b) problems even after prior levels of externalizing and internalizing problems were controlled. These same patterns have been found in cross-cultural longitudinal investigations that find low parent warmth and behavioral over-control in one year are likely to predict greater youth externalizing (Rothenberg et al., 2020a) and internalizing (Rothenberg et al., 2020b) problems the next year in a variety of cultural contexts. However, some studies do find these effects of parenting on subsequent youth externalizing and internalizing problems wane later in adolescence (Lansford et al., 2021a), when adolescent secrecy about activities or other adolescent-driven processes appear to be stronger predictors of adolescent externalizing and internalizing problems (Kapetanovic et al., 2020).

Several longitudinal investigations and systematic reviews have also identified family context variables as especially powerful predictors of subsequent youth externalizing and internalizing problems (Belsky et al., 2020). For instance, quasi-experimental work that examined the effects of family income supplements on adolescent functioning found that those transfers lowered the likelihood of adolescent psychiatric disorder diagnosis years later (Costello et al., 2010). Moreover, family-level socioeconomic status in childhood has been linked to adolescent and young adult mental health years later in several longitudinal cohort studies (Belsky et al., 2020). Similarly, early adverse life events (e.g., death of a family member, divorce) have become so heavily linked with adolescent mental health and functioning that it is now typical to screen for adverse childhood experiences in many medical primary care practices (Belsky et al., 2020). In sum, youth living in family environments marked by adversity and poverty are at greater risk for externalizing and internalizing behavior (Belsky et al., 2020).

Exosystem risk factors

Systematic reviews and cross-cultural investigations have identified the neighborhood one lives in as a risk factor for youth externalizing and internalizing behavior, with youth from more dangerous, high-crime neighborhoods more likely to experience poor mental health (Skinner et al., 2014). Moreover, extant literature indicates that these neighborhood effects may change parenting processes and family contexts that then also impact subsequent adolescent mental health (Leventhal & Dupere, 2019).

Macrosystem risk factors

Cross-cultural investigations have found that the presence of specific beliefs or norms in a cultural context are also associated with adolescent mental health (Lansford et al., 2018b). For instance, some theorists have classified cultures on a continuum from more collectivist in nature (i.e., more focused on establishing group cohesion and social stability) to more individualistic in nature (i.e., more focused on individual achievement and autonomy; Hofstede & Hofstede, 2001). Cross-cultural research provides some evidence that cultures that are more individualistic in nature (such as the United States and Western Europe) experience higher levels of mental health problems than cultures that are collectivist in nature (such as China and Japan; Kessler & Bromet, 2013). Individualism could emerge as a risk factor in these cultural contexts because it weakens the strong social ties, social support, and obligations felt to one's family that might typically protect against the emergence of such symptoms (Lansford et al., 2018a, 2018b). Therefore, the current study also investigates parents' reports of their own individualistic or collectivistic beliefs as potential cultural macrosystem risk factors that predict adolescent mental health problems.

Ordering individual child characteristics, microsystem risk factors, exosystem risk factors, and macrosystem risk factors

As demonstrated in the literature review above, theory-driven work has produced a substantial body of literature implicating each of the child characteristics, and microsystem, macrosystem, and exosystem factors investigated by the present study as risk factors for later adolescent mental health problems. However, existing work has not been able to examine numerous factors from each of these systems simultaneously to determine which might be the most powerful predictors of adolescent mental health (the "information overload" problem referred to above; Tate et al., 2020). Yet, the Bioecological Model does provide clues about the relative power of each of these sets of predictors. Specifically, the Bioecological Model posits that risk factors from systems that are closer to directly contacting a youth might be more powerfully associated with youth mental health (Bronfenbrenner & Morris, 2006). That means that risk factors from systems more proximal to the youth (including the youth's own individual characteristics and the microsystems that they directly interact with) might be more powerful predictors of adolescent mental health than risk factors from systems more distal from the youth (including macrosystem and exosystem factors that primarily impact youth mental health through their effects on microsystems or individual developmental characteristics; Bronfenbrenner & Morris, 2006).

Incorporating the chronosystem

The Bioecological Models' chronosystem suggests that the passage of time might also lead to changes in what most powerfully predicts youth mental health (Bronfenbrenner & Morris, 2006). Specifically, the Bioecological Model posits that as predictors become more distant in developmental time from the behavior that they are predicting, they are often less powerful predictors of that behavior (Cairns et al., 2001). This is because as time passes, other events and actions more proximal to the present also begin to form and shape behavior (Bronfenbrenner & Morris, 2006). For instance, as adolescence progresses, peer relationships become more powerful predictors of adolescent functioning and therefore modestly diminish the predictive power of events from earlier in development (Lansford et al., 2021a).

The Bioecological Model suggests that risk factors that are more proximal to directly contacting a youth, and more proximal in time to the adolescent mental health behavior in question, will be more powerfully predictive of youth mental health. However, because of the computational and replicability difficulties in theory-driven research, little work has been able to examine these hypotheses across each of the Bioecological Model's individual, micro-, macro- and exosystems simultaneously (Dwyer et al., 2018). Data-driven machine learning work has emerged in attempts to address these problems.

Using Machine Learning to Identify the Most Important Predictors of Adolescent Mental Health

Researchers have started to use machine learning to predict adolescent mental health outcomes and identify the most important predictors of adolescent mental health, but the literature is still nascent (Dwyer et al., 2018). One study of 7638 Swedish twins examined 474 predictors collected from ages 9–12 to predict overall problems with mental health at age 15 (Tate et al., 2020). This study identified age 9–12 youth oppositional defiant symptoms, impulsivity symptoms, inattention symptoms, executive dysfunction, emotional symptoms, neighborhood deprivation, peer difficulties, parity, gestational age at birth, and separation anxiety as the top predictors of age 15 mental health problems (Tate et al., 2020). Similarly, a cross-sectional machine learning study predicting the subjective well-being of 10,518 Chinese undergraduates from 298 predictors identified three single-item measures of depression, anxiety, and happiness as the top three most important predictors of subjective well- being in these undergraduates (Zhang et al., 2019). Both of these studies adopted an atheoretical, data-driven approach and were not designed to extensively investigate predictors at each level of the Bioecological Model. For instance, the Tate et al. (2020) study only included one measure that predicted one parenting behavior and the Zhang et al. (2019) study did not included any measure of parenting. However, it is notable that 8 of the top 10 most important predictors of adolescent mental health in the Tate et al. (2020) study, and all 3 of the most important predictors of university student subjective well-being in the Zhang et al. (2019) study could be considered individual characteristics (i.e., mental health behaviors or birth age/order). Thus, the proximity hypothesis made by the Bioecological Model does find preliminary support in both of these studies: predictors more proximal to the child (i.e., individual child characteristics like previous youth mental health) are the most powerful predictors of future youth mental health.

Current Study

Dramatic increases in adolescent mental health problems have given rise to the need for policymakers and interventionists to identify the most important risk factors that predict adolescent mental health problems. One method for doing so, the theory-driven method, identifies a plethora of risk factors but lacks the analytic methodology to identify the most important of these risk factors in a replicable way. A second method for doing so, the data-driven machine learning method, can identify the most important risk factors in a replicable way, but does not have the theoretical framework to easily interpret results. The present study integrates these two methods. Using the Bioecological Model, it makes two hypotheses. First, it hypothesizes that age 10 risk factors that are more proximal to an adolescent (e.g., adolescent mental health history, individual characteristics, parenting, and family context predictors) will be more important predictors of age 13 and age 17 adolescent mental health (i.e., externalizing and internalizing symptoms) than factors that are more distal from the adolescent (e.g., neighborhood quality, individualistic/collectivistic cultural orientations; Hypothesis 1). Second, it hypothesizes that as time goes on, age 10 risk factors will become less powerful predictors of adolescent mental health at age 17, compared to age 13, due to the effects of the chronosystem (Hypothesis 2). Using machine learning methods, the present study tests these hypotheses by investigating which of 79 predictors at the beginning of adolescence (i.e., age 10) are the most powerful predictors of adolescent mental health at age 13 and 17 in 1176 adolescents from 12 ethnocultural groups in 9 countries. In so doing, the current study leverages machine learning's predictive power (Dwyer et al., 2018) while analyzing results through the empirically supported theoretical lens offered by the Bioecological Model (Bronfenbrenner & Morris, 2006).

Methods

Data were collected from the Parenting Across Cultures Project (Lansford et al., 2021b), a longitudinal study of parenting and child mental health that examines the development of over 1000 children in 12 cultural groups from nine nations. This longitudinal investigation has been conducted since children were 8 years old (in 2008) and continues until the present. The measures utilized here come from predictor variables collected when children were age 10 (in approximately 2010) and outcome measure collected when adolescents were ages 13 and 17.

Design

Parenting Across Cultures Project participants who reported on any predictors when children were age 10 were included in the current study. The final study sample consisted of 1176 families with children ($M_{ChildAge} =$ 10.71 years, $SD = .67$; 51% girls). Families were recruited from 12 ethnocultural groups in nine countries including: Shanghai, China ($n = 101$); Medellín, Colombia ($n = 100$); Naples ($n = 95$) and Rome ($n = 99$), Italy; Zarqa, Jordan ($n = 112$); Kisumu, Kenya ($n = 95$); Manila, Philippines ($n = 103$); Trollhättan/Vänersborg, Sweden ($n = 98$); Chiang Mai, Thailand ($n = 101$); and Durham, NC, United States ($n = 100$ White, $n = 92$ Black, $n = 80$ Latinx). These groups were selected because they vary across a number of important dimensions. For example, the countries rank 8^{th}–147^{th} out of 189 countries on the United Nations' Human Development Index, an indicator of a country's health and income status.

Participants were recruited through schools. Response rates varied from 24–100%, primarily because of differences in the schools' roles in recruiting (i.e., some schools took a more active role in recruiting than others). Response rates are unable to be estimated for all sites. In some cases there is no record of the number of students potentially invited to participate versus those who agreed to participate due to the differing ways in which schools informed parents about the study (e.g., letters, email, or verbal announcement). Most parents lived together (82%) and were biological parents (97%); nonresidential and non-biological parents also provided data. Sampling included families from each country's majority ethnic group, except in Kenya where Luo (13% of the population) were sampled, and in the United States, where equal proportions of Black, Latinx, and White families were sampled. SES was sampled in proportions representative of each recruitment area.

Measures were administered in the primary language spoken in the country, following forward- and back-translation. Interviews lasted 2 hours and were conducted after parent consent and child assent were given in participant-chosen locations. At age 10, interviews also included child participation in a battery of observational behavioral tasks that measureds child self-regulation and executive functioning (Duell et al., 2018). Participants were given the choice of completing the measures in writing or orally. Families were given modest monetary compensation for participating or compensated in other ways deemed appropriate by local IRBs.

Measures

Due to the sheer number of predictor variables included in this study (79 in all), measure descriptions are as succinct as possible. Some measures did not emerge as top predictors of age 13 or 17 outcomes. These measures are listed as "Other Measures" within their respective predictor variable categories. To conserve space, extensive description of those measures can be found in the Measures Appendix. However, all measures have demonstrated extensive reliability and validity in past analyses in this sample, and most have also been used previously in large international studies (e.g. Rothenberg et al., 2020a; Rothenberg et al., 2020b). All measures were transformed to z-scores before the machine learning procedure.

Age 10 Child Mental Health Measures

Child externalizing behavior and internalizing behavior

Mothers and fathers completed Achenbach's (1991) Child Behavior Checklist when children were age 10. Children completed the Youth Self Report (Achenbach, 1991) at age 10. Participants were asked to rate how true each item was of their youth (for parents) or themselves (for youth) during the last six months (0 = *not true*, 1 = *somewhat or sometimes true*, 2 = *very or often true*). The *Externalizing Behavior* scale averaged across 33 items (for parent reports) or 30 items (for child reports) captures behaviors such as lying, truancy, vandalism, bullying, drug and alcohol use, disobedience, and physical violence. The *Internalizing Behavior* scale averaged across 31 items (for parent reports) or 29 items (for child reports) measures behaviors and emotions such as loneliness, self-consciousness, nervousness, sadness, and anxiety. All scales were calculated separately for mother, father, and youth reports of youth at age 10. The Achenbach measures are among the most widely used instruments in international research, with translations in over 100 languages and strong, well-documented psychometric properties (e.g., Achenbach & Rescorla, 2006).

Age 10 Other Child Individual Characteristics Measures

Adolescent prosocial behavior

Youth completed a 13-item scale composed of items such as "I try to help others," which was adapted from Pastorelli et al. (1997). Items were rated as 1 = never, 2 = sometimes, or 3 = often. A scale was computed as the average of the 9 prosocial behavior items (the remaining 4 items were distracters).

Adolescent social competence

Mothers and fathers completed a 7-item social competence scale adapted from Pettit et al. (1991) indicating how socially skilled the child was in several kinds of interpersonal interactions (e.g., understanding others' feelings, generating good solutions to interpersonal problems). Items were rated on a 5-point scale from 1 = very poor to 5 = very good. A single scale was computed as the average of the 7 items for both mothers and fathers.

Benthin Risk Perception Scale

This scale, adapted from an instrument developed by Benthin et al. (1993) was designed to measure the extent to which an individual recognizes and evaluates the risks inherent in activities that are potentially dangerous or harmful. Respondents are presented with 8 dangerous activities (e.g., drinking alcohol, having unprotected sex) and are asked to indicate four things for each of these activities: How "scary" the activity is (affective component), how risky the activity is (likelihood component), how much the risks of the activity outweigh its benefits (comparative value component), and how serious the consequences of the activity would be if something "bad" happened as a result (salience component). Each of these ratings is made on a 4-point scale. A single risk perception score was computed by averaging 16 responses (the four evaluation dimensions for four activities).

Modified Iowa Gambling Task

In this task, individuals attempt to earn points by playing or passing cards from four different decks (Bechara et al., 1994). Two of the decks are associated with relatively small gains, but the small gains exceed losses over the course of the task, resulting in a net gain. The other two decks produce larger gains than the first two decks, but in the long run, these decks produce a net loss due to larger losses. In addition, within each type of deck (net gain vs. net loss), there is one deck in which the loss is infrequent but large, and the other deck produces losses that are consistent and small. The ability to choose to pass on bad decks and play on good decks is a measure of decision-making under conditions of uncertainty and risk evaluation. The preference to play the decks with more variance in the outcome is taken as a measure of risk preference. The current study measured both reward sensitivity (based on change in percent of plays on good decks from the first to the last blocks) and cost sensitivity (based on change in percent plays on bad decks from the first to the last blocks).

Tower of London Task

In this task, an individual views a series of three balls on a peg in a start position that must be moved to a pre-specified goal configuration on three other pegs, one of which can support one ball, one of which can support two balls, and one of which can support three balls (Phillips et al., 2001). The subject is instructed to replicate the goal configuration using the smallest number of moves. The number of moves and time required to reach the goal position is measured. Impulse control is measured by the amount of time in seconds to first move on hard problems (i.e., 6 and 7 move problems). Higher scores indicate more impulse control.

Matrix Reasoning

The Matrix Reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999) was used to produce an estimate of nonverbal intellectual ability, which serves as a proxy for IQ in these analyses. Given the variability in language across the research sites, only the Matrix Reasoning subscale t-scores were used in the present analyses.

Child school performance

Mothers and fathers were asked to rate their child's school performance in seven areas (reading, writing, math, social studies, spelling, science, and other). These seven areas were used because they are common to curricula in every country. The questions were adapted from the performance in academic subjects section of the Child Behavior Checklist (CBCL; Achenbach, 1991) which has demonstrated criterion validity. Parents rated whether children were 1 = failing, 2 = below average, 3 = average, or 4 = above average in each area. A single scale was computed as the average of the 7 items to capture total school performance, and subscale scores of performance in each academic area were also computed and included as predictors.

Child gender

Children self-reported their gender as male or female.

Child age

Children self-reported their age in years.

Child pubertal development

At age 10, children completed the Pubertal Development Scale (Petersen et al., 1988), a widely used and well-validated self-report measure of physical development). Five items asked about perceived pubertal changes in skin, height, body hair, and either breast growth and menstruation (for girls) or facial hair growth and voice (for boys). Items were scored on a 0 = has not yet started to 3 = definitely completed scale. Item scores were averaged to create a continuous measure for physical maturation ranging from 0 = puberty has not started to 3 = puberty seems complete.

Other measures

Age 10 Working Memory Accuracy, Verbal Fluency, Balloon Analog, Stoplight, and Stroop tasks were also used to predict adolescent externalizing and internalizing behavior at ages 13 and 17. See Measures Appendix for greater detail.

Age 10 Parenting Predictor Variables

Parent warmth, hostility, neglect, undifferentiated rejection, and behavioral control

All five of these parenting constructs were measured using the appropriate subscales of the Parental Acceptance-Rejection/Control Questionnaire-Short Form (PARQ/Control-SF; Rohner, 2005). Mother, father, and child reports on each construct were collected. Participants rated items for each parent on a modified scale: 1 = never or almost never to 4 = every day.

Parent rules/limit-Setting and parent knowledge solicitation

Both of the constructs were measured via mother, father, and child report. Parent rules/limit-setting and knowledge solicitation were assessed by subscales of the 10-item parental monitoring scale derived from the work of Conger et al. (1994) and Steinberg et al. (1992). To measure parent rules/limit-setting, mothers, fathers, and children answered 5 questions that captured the frequency with which parents impose limits on their child's activities on a 0 = *never* to 3 = *always* scale. To measure parent knowledge solicitation, mothers, fathers, and children answered 5 questions that examined the extent to which parents tried to find out about their child's activities and whom their child spends time with on a 0 = *do not try* to 2 = *try a lot* scale. Both parent rules/limit-setting and parent knowledge solicitation were assessed by asking about the same 5 child activities (e.g., with whom the child spends time, how the child spends his/her free time).

Parent physical punishment

Parent physical punishment was assessed in both mothers and fathers by asking the parent to report the extent to which they physically punish their child on a 0 = never to 5 = almost everyday scale.

Parents' psychological control and autonomy granting

Children rated the extent to which their parents make decisions for them versus let them make their own decisions and how often parents try to control how they think or feel or manipulate them psychologically. This measure yields two subscales, the Psychological Control subscale and the Autonomy Granting subscale, which are both scored on a

0–4 scale with higher scores indicating more control/autonomy granting (Barber, 1996).

Other measures

Age 10 parent discipline strategies and positive parenting strategies were also measured as predictors in the current study (see Measures Appendix).

Age 10 Family Context Predictor Variables

Years of education

Both mother and father total years of education were calculated.

Current employment

Both mothers and fathers reported whether they were currently employed on a 0 = no, 1 = yes item.

Family annual income

The family's annual income over the past year was reported on a 1–10 scale that is scaled such that within each culture, 5 = average income for one's culture, 1 = well below average income for one's culture and 10 = well above average income for one's culture (Rothenberg et al., 2023).

Family obligations

Mothers, fathers, and children reported on this scale, adapted from Fulgini et al. (1999). The Family Obligations Scale asks parents and children to rate a series of statements related to how important it is for children to spend time with their families, help their families, and to respect older members of their families. The first 11 statements ask the parent and child to rate on a scale from 1 = almost never to 5 = almost always how often the child is expected to spend time with and help their family. The other 7 statements ask the parents and children to rate on a scale from 1 = not important to 5 = very important how important certain family practices and relationships are. Items were averaged, with higher scores indicating greater family obligations.

Family Life Events Scale

The Life Events Scale is a series of major life events adapted from Dodge et al. (1994). Parents report whether each of 19 major life events (such as a move, birth of a child, divorce, death of a close family member) has occurred in the last year, indicating either yes or no. In the current study a sum score out of 19 for both father and mother reports of life events over the last year was calculated.

Other measures

How much a family's income changed, the number of individuals in the household, whether parents were divorced or separated, and the importance the family placed on religion were other age 10 predictors measured in this study (see Measures Appendix for greater detail).

Age 10 Neighborhood-Level Predictor Variables

Neighborhood safety

This 7-item scale asked parents and children to rate the safety of the neighborhood in which they live (Skinner et al., 2014). For each reporter, a four-item mean score was used to capture perceptions of neighborhood danger. Higher scores indicate more danger. Mother, father, and child reports were collected.

Age 10 Cultural-Level Predictor Variables

Culture individualism and collectivism

The Individualism and Collectivism portion of the parent interview was adapted from Singelis et al. (1995) and Triandis (1995). Parents rate the importance of different values related to their autonomy and belonging to a social group. Mothers and fathers are asked whether they 1 = strongly disagree, to 4 = strongly agree with a series of 16 statements. Individualistic and collectivistic subscales were scored.

Adolescent Mental Health Outcomes Predicted at Ages 13 and 17

Dichotomous variables indicating "above median" score for child externalizing and/or internalizing behaviors

Adolescent externalizing and internalizing scale scores were computed at ages 13 and 17 using the same items and in the same way as described above in the descriptions of the age 10 predictors (Achenbach, 1991). There was only one difference: When used as outcomes these scales combined mother, father, and adolescent report to create a robust and holistic evaluation of child mental health at these ages. Mother, father, and adolescent reports were combined by averaging the three reports together in line with prior longitudinal, cross-cultural work (Lansford et al., 2018a). To examine if machine learning can be used to examine membership in high-

Table 1 Sample size and number of adolescents for each group, for each target outcome

	Train Partition			Test Partition		
	N	*N* Below [%]	*N* Above [%]	*N*	*N* Below [%]	*N* Above [%]
Age 13 Internalizing Behavior	788	394 [50.0]	394 [50.0]	267	135 [50.6]	132 [49.4]
Age 13 Externalizing Behavior	788	392 [49.7]	396 [50.3]	267	135 [50.6]	132 [49.4]
Age 17 Internalizing Behavior	690	323 [46.8]	367 [53.2]	228	101 [44.3]	127 [55.7]
Age 17 Externalizing Behavior	690	337 [48.8]	353 [51.2]	228	105 [46.1]	123 [53.9]

"*N* Below" indicates the number of adolescents who scored below the median score on externalizing/internalizing behavior measures. "*N* Above" indicates the number of adolescents who scored above the median score on externalizing/internalizing behavior measures

risk-for-mental-health-problems groups, the aforementioned aggregate measures of adolescent externalizing and internalizing behaviors at each of the three ages were z-score centered and then dichotomized to indicate: 1 = adolescents who scored above the median in externalizing problems (or internalizing problems when predicting internalizing behaviors) and 0 = all other adolescents.

Analytic Plan

In this study, a Machine Learning approach was used to identify which of the age 10 measures previously described were more important predictors of the four target outcomes: adolescent internalizing and externalizing behaviors at age 13 and 17. Model performance was assessed using the Matthew Correlation Coefficient (MCC) and Class accuracy ($\%_c$), both computed from the confusion matrix depicted in Appendix Table 1 and described further in the Analytic Plan Appendix. MCC values range from −1 (perfect misclassification) to 1 (perfect classification) and are expected to have values of approximately 0 for random classifiers. Positive values are expected for a model that extracts meaningful information from the dataset and are used to predict the class of each adolescent. MCC is the performance metric used to select the best models (see Analytic Plan Appendix for more detail).

A four-step procedure was followed to build the Machine Learning analytic pipeline. These steps were: a) partitioning of the dataset; b) imputing missing data; c) model optimization and feature selection; d) model selection and performance analysis. This four-step procedure was independently applied for each of the four target age 13 and 17 externalizing and internalizing outcomes.

First, the dataset was randomly split into two partitions: the Train partition (75% of the datapoints) and the Test partition (25% of the datapoints). The Train partition was used to train the data imputation network and in the subsequent steps of the Machine Learning pipeline to build the predictive Machine Learning model, while the Test partition was only used to evaluate the final predictive Machine Learning model (and therefore offers a form of independent validation of the model). Having a separate partition that is only used for evaluation (i.e. the Test partition) allows investigators to assess the generalizability of the Machine Learning model. Only if the patterns extracted by the models from the Train partition are characteristic of the general population will they also be found on the Test partition. The number of adolescents composing each partition and class, for each target outcome, is reported in Table 1.

Second, missingness in both the age 10 predictors and the age 13 and age 17 outcomes was handled. To deal with missingness in age 10 predictors, deep learning procedures were used to impute missing values. Importantly, in accordance with best practice, missing data were only imputed after the test and training data sets were partitioned, to ensure that that estimates of predictive performance were not biased (Kapoor & Narayanan, 2022; Tampu et al., 2022). Missing values were present in 69 of the 79 age 10 predictors. The average age 10 predictor was missing 10.2% of its responses (SD = 8.9%), and missingness ranged from 0.3 to 26.2% across measures. Data imputation was based on a Deep Learning auto-encoder model, inspired by the work of Cheng and colleagues (2020). It appears that the imputation process was successful in accurately imputing missing age 10 predictors (see Data Imputation subsection of the Analytic Plan Appendix and Appendix Figure 1 for further detail).

This data imputation process was not applied to address missingness in the adolescent externalizing and internalizing behavior outcomes at age 13 (where $n = 121$ adolescents; 10% of the sample, were missing data) and age 17 (where $n = 258$ adolescents or 24% of the sample, were missing data). This is because doing so could introduce bias in the data and prevent the objective evaluation of the model performance, since the true value of the missing outcomes is unknown. Therefore, the dataset used to predict age 13 outcomes included data for 1055 adolescents, and the dataset used to predict age 17 outcomes included data for 918 adolescents (see Table 1).

Then, whether those adolescents with versus without missing outcome data at age 13 and 17 systematically differed on any of the 79 age 10 predictors was analyzed (Appendix Table 2). At age 13, adolescents with versus without missing outcome data differed on only 5 of 79 predictors (6.3% of predictors). Compared to adolescents who were missing outcome data at age 13, adolescents with complete outcome data had significantly higher scores on the reward sensitivity measure on the Modified Iowa Gambling Task (Cohen's $d = 0.37$), and neighborhood safety (reported by father; Cohen's $d = 0.28$) and had significantly lower scores on the Balloon Analog Task (Cohen's $d = -0.4$), Matrix Reasoning Task (Cohen's $d = -0.41$), and impulse control measure on the Tower of London Task (Cohen's $d = -0.39$). At age 17, adolescents with versus without missing outcome data differed on only 7 of 79 predictors (8.9% of predictors). Adolescents with complete outcome data at age 17 had significantly higher scores on the following measures: Religious Importance (Cohen's $d = 0.32$), Mother Behavioral Control (Cohen's $d = 0.25$), Knowledge Solicitation (reported by mother; Cohen's $d = 0.45$), Knowledge Solicitation (reported by child; Cohen's $d = 0.28$), and Rules/Limit-setting (reported by mother; Cohen's $d = 0.33$) compared to adolescents who were missing outcome data at age 17 and had significantly lower scores on the Matrix Reasoning Task (Cohen's $d = -0.26$), and on school performance (reported by mother; Cohen's $d = -0.27$). Overall, given that adolescents with versus without outcome data at ages 13 and 17 exhibited few significant differences, it does not appear that adolescents in the age 13 and 17 analytic samples drastically differed from adolescents in the sample as a whole on most age 10 predictors.

The third step in the Machine Learning analytic pipeline was the model optimization and feature selection step, which is called the Data Analysis Plan (DAP). As done in other studies on psychological questionnaires (Gaggero et al., 2021), child development data (Bizzego et al., 2021), and neurophysiological signals (Bizzego et al., 2022b), this study relied on Support Vector Machines (SVMs) with linear kernel as the main Machine Learning model (Cortes & Vapnik, 1995). Although the reported findings mainly refer to the results obtained with SVMs, Random Forests (RF) (Breiman, 2001) were also considered as a comparison model to evaluate the predictive performance. Specifically, ML models were trained to predict whether an adolescent belonged to the *Above* or to the *Below* group, for each of the four target externalizing/internalizing outcomes. Training ML models requires the optimization of the set of predictors and of the model parameters. For the SVM model, the model parameter that was optimized was the regularization coefficient C; for the RF model, the optimized model parameter was the number of trees. The optimization of the model parameter was performed with two nested cross-validation cycles (Bizzego et al., 2022a). In short, the Train partition was randomly split into 5 folds: one fold was used for internal validation (left-out-fold), the other folds were used to train the model. The performances of the trained model were then evaluated on the left-out-fold and the predictive importance of the age 10 predictors was estimated by permutation (see Greater Detail on Model Optimization section of the Analytic Plan Appendix for more information). The procedure was then repeated on all folds, and then from the beginning for 10 times.

The fourth and final step of the Machine Learning procedure included the model selection and performance analysis. First, models selected the optimal set of age 10 predictors, and respective optimal value for the model parameter, that allowed the model to achieve the highest predictive performance in predicting each of the age 13 and 17 externalizing and internalizing behavioral outcomes. Then the final model was trained, with the optimal model parameter and the optimal set of age 10 predictors, using data from the Train partition. Next, the trained final model was applied to data from both the Train and Test partitions and the bootstrapped distribution of the predictive performance was obtained (on both the Train and Test partitions). The C parameter and the model weights were independently optimized on the different age 13 and age 17 externalizing and internalizing behavior target outcomes. A checklist for transparent reporting of machine learning studies (Ali & Ang, 2022) is provided in Appendix Table 3.

A SHAP analysis (Lundberg & Lee, 2017) was also applied to the final trained model, to explain and rank which predictors are most influential in determining the predictive output. SHAP analysis is a method to explain the functioning of machine learning models, based on approaches derived from game theory (shapley values in particular; Lundberg & Lee, 2017). Although SHAP analysis was developed to provide local (i.e., at the datapoint level) explanations about the behavior of the trained model, a global explanation can be also obtained by aggregating the local explanations across the whole dataset. In this study, the SHAP value of each predictor is computed for each datapoint of the Test partition. Then the absolute SHAP values are averaged to obtain a global indication of the relevance of each predictor.

The ranking of the predictors emerging from the SHAP analysis was compared with the averaged rankings obtained from the predictive importance obtained by permutation during the DAP. Specifically, the results section reports the extent to which identified age 10 predictors overlap across the two methods. Additionally, when interpreting which predictors are most important, only age 10 predictors that emerge in both the SHAP and DAP analyses are interpreted. Similarly, when the ranking of each predictor is examined,

Table 2 Median performances of predictive machine learning models as measured by matthew correlation coefficient

			Support Vector Machine			Random Forests		
Age	Behavior	Partition	Median MCC	5% CI	95% CI	Median MCC	5% CI	95% CI
13	Internalizing	Train	0.759	0.685	0.83	0.990	0.981	0.999
		Test	0.443	0.271	0.589	0.453	0.281	0.607
	Externalizing	Train	0.726	0.651	0.793	0.980	0.952	0.998
		Test	0.461	0.29	0.619	0.451	0.286	0.614
17	Internalizing	Train	0.645	0.561	0.733	1.000	0.989	1.000
		Test	0.231	0.043	0.416	0.359	0.170	0.529
	Externalizing	Train	0.802	0.729	0.866	1.000	1.000	1.000
		Test	0.440	0.277	0.62	0.495	0.319	0.660

This table depicts Matthew Correlation Coefficients (MCCs) achieved by the predictive machine learning models across the modeling process, for all target outcomes. CI = Confidence Interval; Train *and* Test = MCCs obtained by the final model on data from the Train and Test partition respectively

only aggregate average predictor rankings across both the SHAP and DAP analyses are reported and considered.

One post-hoc analysis was also conducted. Specifically, the current study statistically investigated whether some target outcomes were better predicted than others (e.g., whether age 13 externalizing behaviors were better predicted than age 17 externalizing behaviors). To achieve this aim, bootstrapped performances achieved on the Test partition (1000 repetitions) were referred to and applied to a one-way analysis of variance (1-ANOVA), followed by post-hoc pairwise comparisons based on T-tests with Bonferroni correction (alpha = 0.05 / 6 pairs of comparisons = 0.083). All analyses were conducted in Python using the scikit-learn package (v0.24.2; Pedregosa et al., 2011), the shap package (Lundberg & Lee, 2017), and the pingouin package (Vallat, 2018)

Results

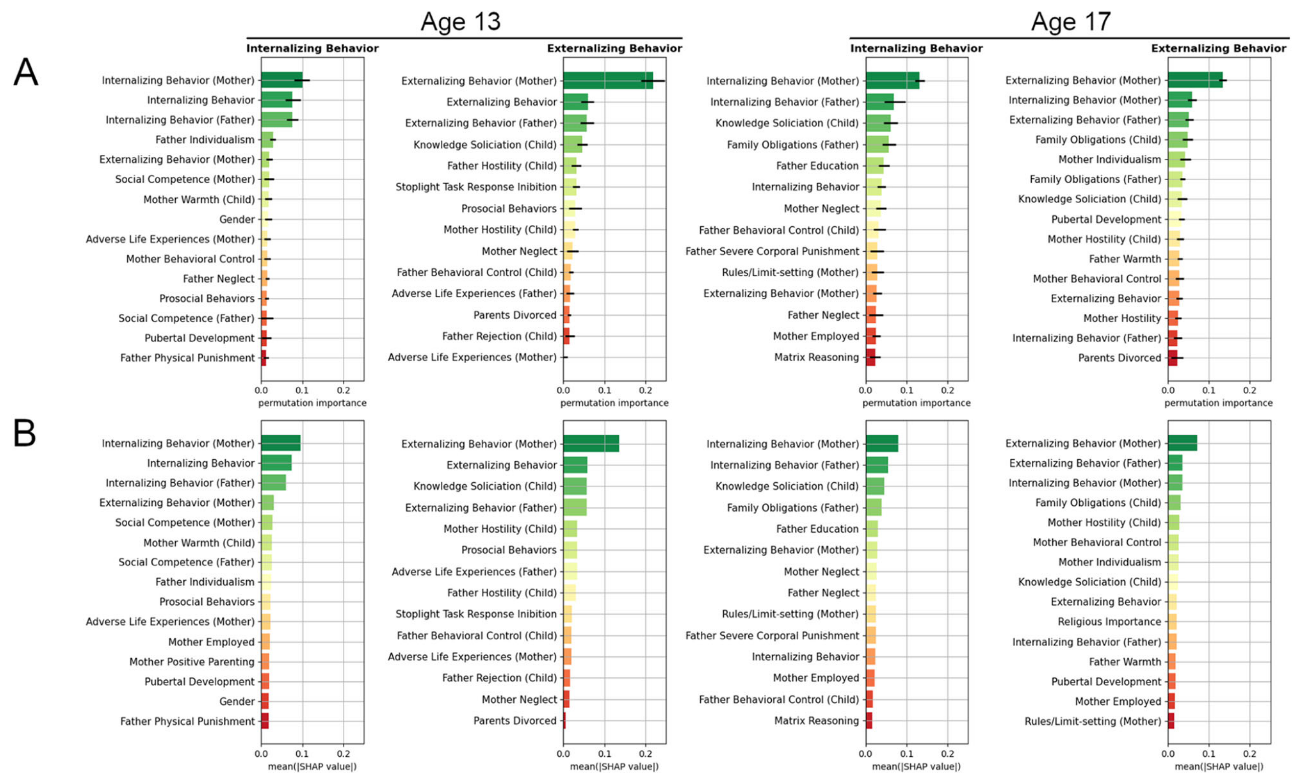
Model Performance

The predictive performances obtained during the Machine Learning procedure in terms of MCC, for each target externalizing/internalizing outcome, for the SVM and Random Forest models, are reported in Table 2.

Overall, it is observed that MCC performance ranges on the Train partition are higher than those on the Test partition, indicating the presence of some overfitting. This effect is much stronger for the RF model, while the confidence intervals of the performance on the Test partitions are comparable between the SVM and the RF. Ultimately, the results of the SVM models (as opposed to the RF models) are presented throughout the rest of the manuscript. This was done because the strong overfitting observed for Random Forest models suggests that the detected patterns, and the consequent model interpretation, may not generalize at the population level. All MCCs were in the positive range, indicating that the Machine Learning predictive models extracted meaningful information from the dataset that was useful to classify each adolescent, and the extracted patterns can be generalized to the entire population of adolescents studied in this sample.

The 1-Way ANOVA examining differences among the Test MCCs for the SVM model, for each of the 4 age 13 and 17 externalizing/internalizing outcomes reported significant results ($F(3, 3996) = 899.7$, $p < 0.001$, partial-eta$^2 = 0.403$). This indicated statistical differences in the Machine Learning model's predictive performance across the age 13 and 17 externalizing/internalizing behavioral outcomes. Post-hoc comparisons with Bonferroni correction (see Appendix Table 4) indicated that the Machine Learning predictive models were best at classifying adolescents according to their age 13 externalizing problems (MCC = 0.461), then approximately equally effective at classifying adolescents according to their age 17 externalizing problems (MCC = 0.440) or age 13 internalizing problems (MCC = 0.443). Model performance was lowest when classifying adolescents according to their age 17 internalizing problems (MCC = 0.231).

Class accuracy for the SVM model is depicted in detail in Appendix Table 5. In general, all the models performed well, and models' abilities to correctly classify adolescents were all well above the chance level (50%). Overall, the pattern of accuracy findings aligns with the pattern of MCC findings reported above. Focusing on the classification of "Above Median" adolescents, the Machine Learning model that used age 10 predictors to predict distal outcomes was most accurate in predicting age 13 internalizing behavior (78.0% of adolescents classified accurately), second most accurate in predicting age 13 externalizing (77.3% of adolescents classified accurately), third most accurate in predicting age 17 externalizing (73.2%) behavior, and least accurate in predicting age 17 internalizing (60.6%) behavior.



Note. Top Row, Row "A": Averages of the permutation importance scores obtained during the Data Analysis Plan. Bottom Row, Row "B": Mean of the absolute SHAP values obtained from the SHAP analysis based on the final model. For Age 13 Externalizing Behavior and Age 17 Internalizing Behavior: the top 14 features are shown, corresponding to the optimal feature set selected during the Data Analysis Plan. Except when explicitly written, predictors refer to the Youth. If the last word is "mother," "father," or "youth," that refers to the reporter reporting on the construct.

Figure 1 The Top 15 Most Important Age-10 Predictors of Each Age 13 and Age 17 Adolescent Outcome

Identifying Most Important Age 10 Predictors of Age 13 & 17 Adolescent Externalizing and Internalizing Behavior

The following subsections discuss the rankings of the age 10 predictors in terms of their importance for predicting age 13 and age 17 outcomes. These rankings were based on the bootstrapped predictive importance computed within the DAP and on the SHAP analysis. The top 15 predictors (out of 79 total predictors) of each of the 4 adolescent outcomes are depicted below in Fig. 1, except in the case of Age 13 Externalizing Behavior and Age 17 Internalizing Behavior, where models only identified 14 predictors as contributing to model prediction.

Age 10 Predictors of Externalizing Behavior

Age 10 predictors of age 13 externalizing behavior

In both the DAP and SHAP analyses, the top four age 10 predictors of age 13 externalizing behavior consisted of three youth mental health predictors (age 10 externalizing behavior as reported by mothers, fathers, and the youth) and one age 10 parenting behavior (parent knowledge solicitation, as reported by the youth). The next ten most important predictors varied slightly in order depending on whether the DAP analyses or SHAP analyses were examined. However, in both analyses, mother and father hostility (as reported by the youth), youth response inhibition as measured by the Stoplight Task, youth-reported prosocial behavior, youth-reported mother neglect, youth-reported father behavioral control, adverse life experiences that a youth experienced (as reported by the father and mother), parental divorce, and youth-reported father rejection emerged among the next 10 most important predictors of Age 13 externalizing behavior. In sum, the DAP and SHAP analyses each identified the exact same set of age 10 predictors as the most important predictors of age 13 adolescent externalizing behavior. These included 3 indicators of youth mental health problems (with an average rank across both DAP and SHAP analyses of 2.17), 2 indicators of other youth individual characteristics (with an average rank across analyses of 7), 6 indicators of parenting behaviors (with an average rank across analyses of 8.33), 3 indicators of family context (with an average rank across analyses of 11.5), and no indicators of neighborhood or cultural effects.

Age 10 predictors of age 17 externalizing behavior

Much like the predictors of age 13 externalizing behaviors, in both the DAP and SHAP analyses, the top four age 10 predictors of age 17 externalizing behavior consisted of 3 youth mental health predictors (age 10 externalizing behavior as reported by mothers and fathers, and age 10 internalizing behaviors as reported by mothers) and 1 age 10 family context variable (family obligations, as reported by the youth). The DAP and SHAP analyses shared 9 of the next 11 most important age 10 predictors of age 17 externalizing behavior. Those shared predictors included: mother self-reported individualism and hostility, and youth-reported parental knowledge solicitation, mother hostility, mother behavioral control, father warmth, youth externalizing behavior, and youth pubertal development. In the DAP analysis, father-reported youth family obligations and whether parents divorced were the other two most important age 10 predictors of age 17 externalizing behavior, whereas in SHAP analysis the importance of religion in the family and whether the mother was employed were the other two most important age 10 predictors. In sum, the DAP and SHAP analyses shared 13 of 15 (87%) age 10 predictors as the most important predictors of age 17 externalizing behavior. These included 5 indicators of youth mental health problems (with an average rank across both DAP and SHAP analyses of 5.8), 1 indicator of other youth individual characteristics (with an average rank across analyses of 10.5), 5 indicators of parenting behaviors (with an average rank across analyses of 9.6), 1 indicator of family context (with an average rank across analyses of 4), 1 indicator of cultural variables (with an average rank of 6 across analyses) and no indicators of neighborhood effects.

Age 10 Predictors of Internalizing Behavior

Age 10 predictors of age 13 internalizing behavior

Continuing the pattern seen in age 10 predictors of externalizing behavior at age 13 and age 17, the top three age 10 predictors of age 13 internalizing behavior in both DAP and SHAP analyses were all measures of age 10 child mental health problems. These three predictors were, in order, mother, youth, and father reports of youth internalizing behavior at age 10.

Once again, as with the other analyses, DAP and SHAP analyses both identified largely the same predictors, but in slightly different orders, when identifying the next 11 most important age 10 predictors of age 13 internalizing behavior. Specifically, both the DAP and SHAP analyses identified 9 of the 11 next most important predictors of age 13 internalizing behaviors including father individualism (father reported), externalizing behavior (mother reported), youth social competence (as reported by the mother and father), youth prosocial behaviors (youth reported), mother warmth (youth reported), youth gender, youth adverse life experiences (mother reported), youth pubertal development, and father physical punishment. Among unshared predictors, the DAP analyses additionally identified mother behavioral control (reported by the youth) and father neglect (reported by the youth) among the most important predictors, while SHAP analyses additionally identified mother employment and mother positive parenting (reported by the youth) among the most important predictors. In sum, the DAP and SHAP analyses shared 13 of 15 (87%) age 10 predictors as the most important predictors of age 13 internalizing behavior. These included 4 indicators of youth mental health problems (with an average rank across both DAP and SHAP analyses of 2.63), 5 indicators of other youth individual characteristics (with an average rank across analyses of 10.5), 2 indicators of parenting behaviors (with an average rank across analyses of 10.75), 1 indicator of family context (with an average rank across analyses of 9.5), 1 indicator of cultural variables (with an average rank of 6 across analyses), and no indicators of neighborhood effects.

Age 10 predictors of age 17 internalizing behavior

As in other models, the top two age 10 predictors of age 13 internalizing behaviors in both DAP and SHAP analyses were measures of youth mental health. They were, in order, mother and father reports of age 10 youth internalizing behavior. DAP and SHAP analyses also each identified the same three next most important predictors of age 17 internalizing behavior, in the same order. They were age 10 parent knowledge solicitation (reported by youth), youth family obligations (reported by father), and father education.

DAP and SHAP analyses also shared the remaining 9 most important age 10 predictors of age 17 internalizing behavior but identified these predictors in slightly different orders. These predictors included: youth Matrix Reasoning score, youth internalizing behavior (youth reported), youth externalizing behavior (mother reported), mother neglect (youth reported), rules/limit-setting (youth reported), mother employment, father behavioral control (youth reported), severe corporal punishment (youth reported), and neglect (youth reported). In sum, the DAP and SHAP analyses shared all 14 of the most important age 10 predictors of age 17 internalizing behavior. These included 4

Table 3 Summary of number of predictors and average rank of predictors of age 13 and 17 adolescent outcomes by age 10 predictor category

Category of Age 10 Predictors	Time Point Wherein Adolescent Externalizing/Internalizing Behavior is Being Predicted								
	Age 13			Age 17			Total		
	% of All Predictors in Category That Were Top 15 Predictors	# of Predictors in Top 15	Average Rank of Predictors in Top 15 Predictors	% of All Predictors in Category That Were Top 15 Predictors	# of Predictors in the Top 15	Average Rank of Predictors in Top 15	% of All Predictors in Category That Were Top 15 Predictors	# of Predictors in the Top 15	Average Rank of Predictors in Top 15
Youth Mental Health Problems	39%	7	2.43	50%	9	5.44	44%	16	3.94
Other Youth Individual Characteristics	23%	7	9.21	7%	2	12.25	15%	9	9.89
Parenting Behaviors	12%	8	8.93	16%	11	8.86	14%	19	8.89
Family Context	14%	4	11	14%	4	6.38	14%	8	8.69
Neighborhood Effects	0%	0	N/A	0%	0	0	0%	0	N/A
Cultural Variables	12.5%	1	6	12.5%	1	6	12.5%	2	6

Only predictors that made the Top 15 lists in both the DAP and SHAP analysis (see Fig. 1) were included in these calculations. Additionally, for Age 13 externalizing behavior and Age 17 internalizing behavior, only 14 age 10 predictors (instead of 15) emerged as meaningfully impactful predictors. However, the label "Top 15" is used for the sake of parsimony and simplicity

indicators of youth mental health problems (with an average rank across both DAP and SHAP analyses of 5), 1 indicator of other youth individual characteristics (with an average rank across analyses of 14), 6 indicators of parenting behaviors (with an average rank across analyses of 8.25), 3 indicators of family context (with an average rank across analyses of 7.17), and no indicators of neighborhood or cultural effects.

Summary of Support for Aim 1 and 2 Hypotheses

Hypothesis 1 posited that age 10 risk factors that resided in bioecological systems that were more proximal to the adolescent would be more important predictors of age 13 and age 17 mental health than age 10 risk factors that resided in bioecological systems that were more distal from the adolescent. That is, predictors that were part of an adolescent (i.e., their history of mental health problems and individual characteristics) would be the most powerful predictors of age 13 and age 17 mental health, followed by microsystem predictors that the adolescent has direct contact with (i.e., parenting behaviors and family context), followed by macrosystem and exosystem predictors that affect adolescents mainly through changes to the microsystem (i.e., neighborhood effects and cultural variables). A summary of the number of top 15 age 10 predictors and average rank of these predictors by category can be found in Table 3.

As can be seen in Table 3, this hypothesis was generally supported. Age 10 child mental health problems did appear to be the most powerful age 10 predictors of ages 13 and 17 mental health (with 44% of all child mental health predictors landing on top 15 lists for a total of 16 predictors and an average rank of 3.94 of those predictors), Then, family context variables (14% of all family context predictors landing on top 15 lists, for a total of 8 predictors and an average rank of 8.69), parenting behaviors (14% of all parenting predictors landing on top 15 lists, for a total of 19 predictors and an average rank of 8.89), and other youth individual characteristics (15% of all youth individual characteristics landing on top 15 lists, for a total of 9 predictors and an average rank of 9.89) landed in an interchangeable tier as the next most important age 10 predictors of age 13 and 17 adolescent mental health. Finally, cultural-level variables (which only made it into 2 top 15 predictor lists), and neighborhood effects (which did not make it on any list) were the least important predictors of age 13 and 17 adolescent mental health.

The order of these categories was somewhat, but not completely, stable over time (Table 3). Age 10 child mental health problems appeared to be the most important predictors, and age 10 neighborhood and cultural variables appeared to be the least important predictors, of both age 13 and age 17 externalizing and internalizing problems. However, the middle tier of other youth individual characteristics, parenting behaviors, and family context age 10 predictors varied slightly in their predictive importance from age 13 and 17. At age 13, parenting variables appeared to be the most important predictors of age 13 adolescent externalizing and internalizing behavior (with 8 predictors on Top 15 lists and an average rank of 8.93 for those predictors), followed by other individual youth characteristics (7 predictors on top 15 lists, average rank of 9.21), and then family context variables (4 predictors on top 15 lists, average rank of 11). At age 17, family context variables appeared to be the most important predictors of age 17 adolescent externalizing and internalizing problems (with 4 predictors landing on top 15 lists, average rank of 6.38), followed by parenting variables (11 predictors landing on top 15 lists, average rank of 8.86), and then other individual youth characteristics (2 predictors on top 15 lists, average rank of 12.25).

In sum, as expected, individual characteristics (i.e., youth mental health problems at age 10) appeared to be the most important predictor of age 13 and 17 adolescent mental health, followed by indicators of the microsystem (parenting behaviors and finally context variables), and then finally indicators of the macro and exosystem (cultural variables and neighborhood variables). The one exception to the general support found for Hypothesis 1 was the result that parenting and family context variables (microsystem indicators) appeared to be just as important to predicting later adolescent mental health as other adolescent individual characteristics (other than age 10 adolescent mental health problems). Hypothesis 1 predicted that other individual characteristics may have been slightly more important in predicting adolescent mental health outcomes, because they reside within the individual, and are thus slightly more proximal to the individual than parenting and family context microsystem indicators.

Hypothesis 2 posited that the total predictive power of the age 10 predictors would diminish from age 13 to age 17. This hypothesis was partially supported. As described in the Model Performance section, machine learning models did predict an age 13 mental health problem (externalizing problems) significantly better than all other models and did predict an age 17 mental health problem (internalizing problems) significantly worse than all other models. These results support Hypothesis 2. However, machine learning models also were equally effective in predicting age 13 internalizing problems and age 17 externalizing problems. These results run contrary to Hypothesis 2's expectations. In other words, instead of finding that the predictive power of age 10 predictors diminished over time for all age 17 models, this study found that such predictive power diminished over time only in some age 17 models (i.e., those predicting internalizing behavior).

Sensitivity Analyses Predicting Adolescent Membership In Highest or Lowest Quartiles

It may be that different predictors predict whether an adolescent was experiencing extremely high levels of mental health problems, compared to whether an adolescent was merely "above or below the median" on mental health problems. Therefore, the exact same analytic procedure described above was used to examine which age 10 predictors were most powerful in predicting which adolescents were in the upper quartile of age 13 and 17 externalizing/internalizing behavior. These sensitivity analyses revealed that essentially the same set of age 10 predictors that predicted adolescent membership in the "above median" groups emerged as the most powerful predictors of adolescent membership in the upper quartile of externalizing/internalizing behavior. However, in these sensitivity analyses, there was some evidence of model overfit which threatened results generalizability. Therefore, the median-split analyses are presented as the main findings, and the "upper quartile" analyses as sensitivity analyses. Sensitivity analysis results are available from the first author upon request.

Sensitivity Analyses Predicting Age 13 and 17 Behaviors Removing Age 10 Externalizing/Internalizing Behavior

Results indicated that age 10 externalizing and internalizing behavior were consistently the strongest predictors of age 13 and 17 externalizing and internalizing behavior. It was unknown how well models performed, and how stable the other age 10 predictors of age 13 and 17 externalizing and internalizing behavior were, if these powerful age 10 externalizing and internalizing behavioral predictors were removed from the model. Therefore, sensitivity analyses were performed where age 10 internalizing/externalizing behaviors predictors were removed from the dataset, and a new full training procedure was executed. Full results are presented in the Results Without Age 10 Internalizing/Externalizing Behavior Predictors section of the Appendix as well as Appendix Table 6 and Appendix Figure 2. In brief, these sensitivity analyses revealed that when age 10 externalizing and internalizing predictors were excluded from the predictive models, predictive model performance did decrease, but not to an extent that they were statistically significant different from the models that included these age 10 predictors. The sole exception to this was the model predicting age 13 externalizing behavior, which did perform significantly worse when age 10 externalizing and internalizing predictors were removed from the predictor set.

Moreover, even models that exclude age 10 externalizing and internalizing predictors still provided information that meaningfully predicted adolescent internalizing and externalizing behaviors at ages 13 and 17. Additionally, these sensitivity analyses appeared to indicate that once age 10 externalizing and internalizing predictors were removed, the other predictors that predict age 13 and 17 externalizing and internalizing behavior remain relatively stable. Specifically, of the 41 non-age 10 externalizing/internalizing predictors that landed among the most important predictors in original analyses (Fig. 1), 28 (68.29%) of these predictors remained among the most important predictors in the sensitivity analyses (Range: 60–80% across ages 13–17 externalizing and internalizing behavior; Appendix Figure 2). In sum, these sensitivity analyses indicate that though including age 10 externalizing and internalizing behaviors as predictors is important for improving model performance, models still perform well and offer useful predictive value when those predictors are excluded. Moreover, most of the most powerful predictors of age 13 and 17 externalizing and internalizing behavior remain the same across both sets of models.

Discussion

Identifying the most important risk factors that predict the emergence of mental health outcomes in adolescence is especially important, given that globally, the average age of onset for mental disorders is 14.5 years (Solmi et al., 2021) and adolescent mental health problems are increasing (Benton et al., 2021). One method for doing so, the theory-driven method, identifies a plethora of risk factors but lacks the analytic methodology to identify the most important of these risk factors in a replicable way (Dwyer et al., 2018). A second method for doing so, the data-driven machine learning method, can identify the most important risk factors in a replicable way, but does not have the theoretical framework to easily interpret results (Dwyer et al., 2018). The present study contributes to existing literature by demonstrating how these two methods can be integrated to identify the most important age 10 risk factors for predicting adolescent mental health over time. Using the Bioecological Model within a machine learning framework, the present study finds that the most important age 10 risk factors for predicting mid (i.e., age 13) and late (i.e., age 17) adolescent mental health are mental health problems and aspects of the youth's microsystem (i.e., parenting behaviors and the family context). It also identifies that preadolescent risk factors appear to diminish in their predictive power over time when it comes to predicting late adolescent internalizing, but not externalizing, behavior.

Examining Support for Hypothesis 1

Overall, findings demonstrate support for Hypothesis 1. In this sample, the Bioecological Model's conceptualization of the child developing within a multi-layered bioecological system is supported (Bronfenbrenner & Morris, 2006). Though invoked frequently in literature explaining the etiology of adolescent mental health problems, the Bioecological Model's distinct subsystems (i.e., the individual, microsystem, exosystem, and macrosystem) are virtually never examined as simultaneous predictors of adolescent mental health. Machine learning analytic methods allowed the existing study to advance existing literature by doing so (Dwyer et al., 2018).

As the Bioecological Model predicts, it appears that a youth's individual characteristics (particularly their mental health problems) at age 10 serve as the most powerful predictors of a youth's mental health problems throughout adolescence. Indeed, age 10 youth internalizing and externalizing problems are among the top two predictors in every model predicting age 13 and 17 youth internalizing and externalizing problems (Fig. 1). These results align with meta-analyses that demonstrate past externalizing (Pinquart, 2017a) and internalizing (Pinquart, 2017b) mental health problems are the strongest predictors of future mental health problems. However, the present analyses expand on these existing meta-analyses by demonstrating that these prospective associations persist even across 7-year intervals during adolescence and remain strong even when competing with 70+ other predictor variables.

The Bioecological Model predicts that after an individual's own history and characteristics, the next set of predictors that might be most powerful in predicting youth mental health problems are indicators of the microsystem (i.e., those aspects of the child's environment that have direct contact with the child). This was also demonstrated in the current findings; family context variables and parenting behaviors ranked as the two next most important categories of age 10 predictors that predicted adolescent mental health at ages 13 and 17. Of all the parenting variables, variables that indicate parent rejection (e.g., hostility, neglect, lack of warmth) were especially pernicious predictors of later adolescent mental health problems. Some measure of parent rejection appears among the top 10 most important predictors across all adolescent externalizing and internalizing outcomes measured here. This finding aligns with work that asserts parent rejection is universally damaging across cultures, and persistent across time (Rohner, 2005).

Interestingly, no single family context variable seemed to predict adolescent mental health problems at both ages 13 and ages 17. Instead, different constellations of family context variables appeared to predict specific adolescent mental health problems at specific times. For instance, age 10 adverse life experiences predicted age 13 internalizing problems across both DAP and SHAP models; age 10 adverse life experiences and parental divorce predicted age 13 externalizing problems across both models; age 10 family obligations, father education, and maternal employment predicted age 17 internalizing problems; and age 10 family obligations predicted age 17 externalizing problems. It may be that, though family contexts as a whole matter in predicting future adolescent mental health, the exact components of those contexts that matter most for particular mental health problems vary across ontogeny (Bornstein et al., 2021).

Finally, the Bioecological Model also posits that components of the exosystem (e.g., the neighborhoods and communities adolescents reside in) and macrosystem (e.g., the values of cultures adolescents reside in) might be less directly impactful on adolescent mental health, because their effects are largely mediated through more proximal microsystem components (e.g., cultural norms influence parenting which influences child mental health; Lansford et al., 2018b). The current findings support this assertion; exosystem (i.e., neighborhood safety) predictors were never among the top 15 predictors of later adolescent mental health and macrosystem predictors (i.e., individualism) rarely were. These results do not dismiss macrosystem and exosystem predictors as unimportant to adolescent mental health; they just indicate that their power as predictors is lower than (and possibly mediated through) microsystem and individual characteristic predictors (Bronfenbrenner & Morris, 2006).

The exception to the general support for Hypothesis 1 in the current study was the finding that parenting and family context variables (microsystem indicators) appeared to be just as important to predicting later adolescent mental health as other adolescent individual characteristics (other than age 10 adolescent mental health problems). Indeed, the three categories of predictors each had between 14–15% of predictors in that category that landed on top-15 predictor lists, and the average rank of predictors in all three categories ranged between 8.69–9.89 across all time points (Table 3). It was expected that adolescent individual characteristics (e.g., gender, intelligence, emotion regulation) would be slightly more important predictors of adolescent mental health compared to parenting and family context, because these individual characteristics are relatively stable, internal states through which most adolescent experience is filtered (Bronfenbrenner & Morris, 2006). Instead, the current results suggest a slight reorientation of the Bioecological Model when applied to the study of adolescent mental health problems across ontogeny. Specifically, perhaps instead of conceptualizing an *individual* adolescent as being embedded in a microsystem that consists of parenting and family context, it is more useful to conceptualize *a mental health history or trajectory* as something embedded in a microsystem that consists of parenting, family context, and other individual characteristics (intelligence, emotion regulation, gender) that are each direct proximal predictors of the *unfolding of that mental health history over adolescence* (Cairns et al., 2001).

Examining Support for Hypothesis 2

Hypothesis 2 was partially supported; it appears the chronosystem worked in different ways for different adolescent mental health problems. Instead of finding that the predictive power of age 10 predictors diminished over time, such predictive power diminished over time only in some age 17 models (i.e., those predicting internalizing behavior) but not others (i.e., those predicting externalizing behavior). Indeed, age 17 externalizing behavior models correctly classified 73.2% of adolescents above the median in externalizing behavior, a number only slightly smaller than the 78.0% and 77.3% correct classification rates of age 13 externalizing and internalizing models, respectively. The age 17 internalizing behavior model lagged behind the other three, with only a 60.6% correct classification rate. The differential model performance in predicting age 17 externalizing and internalizing problems might have to do with the differential timing of onset of these types of problems (Solmi et al., 2021). A recent meta-analysis of over 700,000 individuals worldwide found that the peak age of onset for many externalizing disorders is 9.5 years old, whereas the peak age of onset for many internalizing disorders (especially depression) is 19.5 years old (Solmi et al., 2021). Therefore, age 10 predictors might more stably predict adolescent externalizing problems across time because the conditions that bring about such problems might already be in place at age 10 (Solmi et al., 2021). In contrast, age 10 predictors might be less accurate in predicting adolescent internalizing problems across time because the conditions that bring about the most severe of those problems do not coalesce until much later in adolescence. This hypothesis needs further testing in future work.

Additionally, the machine learning algorithms using age 10 predictors were able to accurately estimate whether someone was above average in their externalizing or internalizing behavior about 7 out of 10 times, depending on which age and mental health category was examined. This prediction is significantly better than chance. Systematic reviews of widely used mental health checklist measures reveal that they are usually able to correctly identify children with diagnosable mental health problems with somewhere between 60–80% accuracy (Lavigne et al., 2016). Therefore, an imperfect but potentially useful comparison to draw is that the machine learning models examined in this study are as accurate in predicting which adolescents might be above or below average in their mental health problems in 7 years as mental health screeners are at identifying adolescents who have mental health disorders in the moment. Such long-lasting accuracy may be clinically useful (Solmi et al., 2021).

Implications for Future Integration of Theory-Driven and Data-Driven Models of Adolescent Development

In sum, the general alignment between the Bioecological Model's theory-driven work (Bronfenbrenner & Morris, 2006) and atheoretical, data-driven machine learning models found in the current results is encouraging for the field of developmental psychopathology. It suggests that the predominant theoretical paradigms in developmental psychopathology hold up well even when examined via atheoretical machine learning methods. In other words, the current machine learning results serve as an independent verification of the robustness of the Bioecological Model of mental health and lend confidence in the continued use of this model. Given that the replication crisis has cast doubt on the utility of a priori theory and the inferential statistics that accompany theories, it is encouraging to see evidence of this verification (Dwyer et al., 2018).

These results also have important implications for the application of machine-learning algorithms to adolescent science. They demonstrate that machine learning algorithms are indeed effective at distilling a plethora of risk factors down to the most important predictors of subsequent adolescent health (Dwyer et al., 2018). They also demonstrate that, when paired with theory, such results are more readily interpretable. Instead of trying to make sense of a mass of important predictors without any theoretical lens through which to interpret such results, the current study results use the Bioecological Model to determine that individual mental health histories and microsystem indicators appear to be the most important predictors of subsequent adolescent mental health. This novel integration of theory-driven and data-driven approaches overcomes the weaknesses in each approach to provide new insights into the development of adolescent mental health problems. In this study, this integration resulted in predictive models that generally give interventionists a 7 in 10 chance of accurately predicting which adolescents might be experiencing elevated mental health problems 3–7 years from the present.

Implications For Screening and Treatment

The current study has three implications for screening and treatment. First, this study adds to myriad evidence that suggests screening for child and adolescent mental health problems regularly is a very good idea (Solmi et al., 2021). However, this study provides a new justification for such screening: It is likely to not only be accurate now, but also useful in predicting adolescent mental health outcomes years down the road. For that reason, investing in creating systematic universal child and adolescent mental health screening tools in schools, primary care settings, or both, is important (Lavigne et al., 2016). Second, this study suggests that if one had to select a single measure to screen for future adolescent mental health problems, choosing a measure of current youth mental health problems would be wise. The axiom "the best predictor of future behavior is past behavior" holds strong here. The top 2 predictors of adolescent externalizing and internalizing problems at every single time point were all measures of past child mental health problems. Third, this study also suggests that if one can only use one reporter to screen for adolescent mental health problems, that reporter should probably be the mother. Mother reports of each externalizing and internalizing problem at age 10 were the top predictors of adolescent mental health problems at ages 13 and 17.

Limitations & Future Directions

The current study has several limitations. First, the study included relatively small sample sizes within each culture which limits the number of predictors that could be used in the machine learning models and the number of people that could be withheld from the machine learning training sample to serve as the test sample. Future studies with larger samples could test the relative importance of even more predictors in training sets that would be even more likely to generalize. Second, the current study did not include measures of peer affiliation or influence among age 10 predictors. However, peer processes undoubtedly loom large as a set of variables that can predict adolescent mental health development (Lansford et al., 2021a). Future work would do well to include these peer processes as predictor variables in machine learning models. Third, the current study only evaluated individual characteristics and microsystem, exosystem, and macrosystem variables as *direct* predictors of adolescent mental health outcomes. However, Bioecological Theory posits that macrosystem and exosystem variables primarily impact individual development through their influence on microsystem variables (Bronfenbrenner &

Morris, 2006). The current study was unable to examine these mediating processes. Future studies that combine machine learning algorithms and inferential statistical mediation models could do so. Finally, it should be noted that the current study also examined fewer predictors in the macrosystem and exosystem than the microsystem. In addition to examining neighborhood safety and individualism and collectivism, future studies could investigate the predictive effects of other macrosystem and exosystem variables, including variables like government policies, mental health access, and cultural norms around mental health stigma on adolescent mental health development.

Conclusion

Theory-driven research has identified a plethora of risk factors that predict adolescent mental health problems but has difficulty distilling and replicating these findings. Data-driven machine learning methods can distill risk factors to identify which are the most important in predicting adolescent mental health, and can replicate these findings readily, but have difficulty interpreting findings because they are largely atheoretical. The present study contributes to existing literature by demonstrating how these two methods can be integrated to identify the most important early risk factors in predicting adolescent mental health over time. Using the Bioecological Model within a machine learning framework, the present study finds that the most important risk factors for predicting mid (i.e., age 13) and late (i.e., age 17) adolescent mental health are preadolescent mental health problems and aspects of the youth's microsystem (i.e., parenting behaviors and the family context). It also identifies that preadolescent risk factors appear to diminish in their predictive power over time when it comes to predicting late adolescent internalizing, but not externalizing, behavior. This combination of theoretical and machine-learning models can accurately predict which adolescents demonstrate above average mental health difficulties 3–7 years later in approximately 7 out 10 adolescents. The current study demonstrates the power of machine learning to evaluate developmental and clinical theory and predict future adolescent mental health problems.

Acknowledgements This research has been funded by NICHD grant RO1-HD054805 and Fogarty International Center grant RO3-TW008141. This research also was supported by the Intramural Research Program of the NIH/NICHD, USA, and the Centre for the Evaluation of Development Policies (EDePO) at the Institute for Fiscal Studies (IFS), London, UK, funded by the European Research Council (ERC) under the Horizon 2020 research and innovation programme (grant agreement No 695300-HKADeC-ERC-2015-AdG). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or NICHD. We are indebted to the now-decades-long participation of the Parenting Across Cultures project families in our work. None of what we do would be possible without our amazing families.

Authors' contributions W.A.R. conceived of the study, participated in its design and coordination and drafted the manuscript; A.B. developed the analytic plan, performed study analyses and participated in the drafting of the manuscript; GE developed the analytic plan and participated in the drafting of the manuscript; J.E.L. designed study data collection, participated in study design and coordination, obtained funding for the study, and drafted the manuscript. S.M.A.H. designed the study, coordinated study data collection at each site, and revised the manuscript; D.B. designed the study, coordinated study data collection at each site, and revised the manuscript; M.H.B. designed the study, coordinated study data collection at each site, and revised the manuscript; L.C. designed the study, coordinated study data collection at each site, and revised the manuscript; K.D.D. designed the study, coordinated study data collection at each site, and revised the manuscript; L.D. designed the study, coordinated study data collection at each site, and revised the manuscript; K.A.D. designed the study, coordinated study data collection at each site, and revised the manuscript; S.G. designed the study, coordinated study data collection at each site, and revised the manuscript; Q.L. designed the study, coordinated study data collection at each site, and revised the manuscript; Q.L. designed the study, coordinated study data collection at each site, and revised the manuscript; P.O. designed the study, coordinated study data collection at each site, and revised the manuscript; C.P. designed the study, coordinated study data collection at each site, and revised the manuscript; A.T.S. designed the study, coordinated study data collection at each site, and revised the manuscript; E.S. designed the study, coordinated study data collection at each site, and revised the manuscript; S.T. designed the study, coordinated study data collection at each site, and revised the manuscript; L.S. designed the study, coordinated study data collection at each site, and revised the manuscript; L.M.U.T. designed the study, coordinated study data collection at each site, and revised the manuscript; S.Y. designed the study, coordinated study data collection at each site, and revised the manuscript; L.P.A. designed the study, coordinated study data collection at each site, and revised the manuscript. All authors read and approved the final manuscript.

Funding This research has been funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development grants RO1-HD054805 and P2CHD065563 and Fogarty International Center grant RO3-TW008141. This research also was supported by National Institute on Drug Abuse (NIDA) Grant P30 DA023026, the Intramural Research Program of the NIH/NICHD, USA, and an International Research Fellowship at the Institute for Fiscal Studies (IFS), London, UK, funded by the European Research Council (ERC) under the Horizon 2020 research and innovation program (grant agreement No. 695300-HKADeC-ERC-2015-AdG).

Data sharing declaration The datasets generated and/or analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committees (Duke University

Institutional Review Board Protocol 2032 and ethics committees at universities in all nine participating countries) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2006). Developmental issues in assessment, taxonomy, and diagnosis of psychopathology: Life span and multicultural perspectives. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Theory and method* (pp. 139–180). John Wiley & Sons, Inc.

Ali, F., & Ang, R. P. (2022). Predicting how well adolescents get along with peers and teachers: A machine learning approach. *Journal of Youth and Adolescence*, *51*(7), 1241–1256. https://doi.org/10.1007/s10964-022-01605-5.

Barber, B. K. (1996). Parental psychological control: Revisiting a neglected construct. *Child Development*, *67*(6), 3296–3319. https://doi.org/10.2307/1131780.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15. https://doi.org/10.1016/0010-0277(94)90018-3.

Belsky, J., Caspi, A., Moffitt, T. E., & Poulton, R. (2020). *The origins of you: How childhood shapes later life*. Harvard University Press.

Benthin, A., Slovic, P., & Severson, H. (1993). A psychometric study of adolescent risk perception. *Journal of Adolescence*, *16*(2), 153–168. https://doi.org/10.1006/jado.1993.1014.

Benton, T. D., Boyd, R. C., & Njoroge, W. F. (2021). Addressing the global crisis of child and adolescent mental health. *JAMA Pediatrics*, *175*(11), 1108 https://doi.org/10.1001/jamapediatrics.2021.2479.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Bizzego, A., Gabrieli, G., Bornstein, M. H., Deater-Deckard, K., Lansford, J. E., Bradley, R. H., & Esposito, G. (2021). Predictors of contemporary under-5 child mortality in low-and middle-income countries: A machine learning approach. *International Journal of Environmental Research and Public Health*, *18*(3), 1315 https://doi.org/10.3390/ijerph18031315.

Bizzego, A., Gabrieli, G., Lim, M., Rothenberg, W.A., Bornstein, M.H., & Esposito, G. (2022a). Predictors of early childhood development. In M. H. Bornstein, W. A. Rothenberg, J. E. Lansford, R. H. Bradley, K. Deater-Deckard, S. Zietz, D. L. Putnick, A. Bizzegio, & G. Esposito (Eds.), *Parenting and child development in low- and middle-income countries* (pp. 210–239). Routledge.

Bizzego, A., Neoh, M., Gabrieli, G., & Esposito, G. (2022b). A machine learning perspective on fNIRS signal quality control approaches. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *30*, 2292–2300. https://doi.org/10.1109/TNSRE.2022.3198110.

Bloemen, A. J. P., Oldehinkel, A. J., Laceulle, O. M., Ormel, J., Rommelse, N. N. J., & Hartman, C. A. (2018). The association between executive functioning and psychopathology: General or specific? *Psychological Medicine*, *48*(11), 1787–1794. https://doi.org/10.1017/S0033291717003269.

Bornstein, M. H., & Rothenberg, W. A. (2022). The UNICEF multiple indicator cluster surveys and early childhood development index. In M. H. Bornstein, W. A. Rothenberg, J. E. Lansford, R. H. Bradley, K. Deater-Deckard, S. Zietz, D. L. Putnick, A. Bizzegio, & G. Esposito (Eds.), *Parenting and child development in low- and middle-income countries* (pp. 240–275). Routledge.

Bornstein, M. H., Rothenberg, W. A., & Lansford, J. E. (2021). Change over time in four domains of parenting in diverse international contexts: Specificity and universality, country and culture, determinants, strengths and limitations, future directions and implications. In J. E. Lansford, W. A. Rothenberg, & M. H. Bornstein (Eds.), *Parenting across cultures across development: Parenting from childhood to adolescence in nine countries* (pp. 227–263). Routledge.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Theoretical models of human development., Vol. 1*, 6th ed. (pp. 793–828). John Wiley & Sons Inc.

Cairns, R. B., Elder, G. H., & Costello, E. J. (2001). *Developmental science*. Cambridge University Press.

Cheng, C. Y., Tseng, W. L., Chang, C. F., Chang, C. H., & Gau, S. S. F. (2020). A deep learning approach for missing data imputation of rating scales assessing attention-deficit hyperactivity disorder. *Frontiers in Psychiatry*, *11*, 673 https://doi.org/10.3389/fpsyt.2020.00673.

Conger, R. D., Ge, X., Elder, G. H., Lorenz, F. O., & Simons, R. L. (1994). Economic stress, coercive family process, and developmental problems of adolescents. *Child Development*, *65*(2), 541–561. https://doi.org/10.2307/1131401.

Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*, 273–297.

Costello, E. J., Erkanli, A., Copeland, W., & Angold, A. (2010). Association of family income supplements in adolescence with development of psychiatric and substance use disorders in adulthood among an American Indian population. *JAMA: Journal of the American Medical Association*, *303*(19), 1954–1960. https://doi.org/10.1001/jama.2010.621.

Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. *Child Development*, *65*(2), 649–665. https://doi.org/10.2307/1131407.

Duell, N., Steinberg, L., Icenogle, G., Chein, J., Chaudhary, N., Di Giunta, L., Dodge, K. A., Fanti, K. A., Lansford, J. E., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Uribe Tirado, L. M., Alampay, L. P., Al-Hassan, S. M., Takash, H. M. S., Bacchini, D., & Chang, L. (2018). Age patterns in risk taking across the world. *Journal of Youth and Adolescence*, *47*(5), 1052–1072. https://doi.org/10.1007/s10964-017-0752-y.

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology*, *14*, 98–118. https://doi.org/10.1146/annurev-clinpsy-032816-045037.

Fuligni, A. J., Tseng, V., & Lam, M. (1999). Attitudes toward family obligations among American adolescents with Asian, Latin American, and European backgrounds. *Child Development*, *70*(4), 1030–1044. https://doi.org/10.1111/1467-8624.00075.

Gaggero, G., Bizzego, A., Dellantonio, S., Pastore, L., Lim, M. & Esposito, G. (2021). Clarifying the relationship between alexithymia and subjective interoception. *PLoS ONE*, *16*(12), e0261126. https://doi.org/10.1371/journal.pone.0261126.

Hofstede, G. H., & Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage.

Icenogle, G., Steinberg, L., Duell, N., Chein, J., Chang, L., Chaudhary, N., Di Giunta, L., Dodge, K. A., Fanti, K. A., Lansford, J. E., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Uribe Tirado, L. M., Alampay, L. P., Al-Hassan, S. M., Takash, H. M. S., & Bacchini, D. (2019). Adolescents' cognitive capacity reaches adult levels prior to their psychosocial maturity: Evidence for a “maturity gap” in a multinational, cross-sectional sample. *Law and Human Behavior*, *43*(1), 69–85. https://doi.org/10.1037/lhb0000315.supp. (Supplemental).

Kapetanovic, S., Rothenberg, W. A., Lansford, J. E., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Gurdal, S., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Uribe Tirado, L. M., Yotanyamaneewong, S., Peña Alampay, L., & Bacchini, D. (2020). Cross-cultural examination of links between parent–adolescent communication and adolescent psychological problems in 12 cultural groups. *Journal of Youth and Adolescence*, *49*(6), 1225–1244. https://doi.org/10.1007/s10964-020-01212-2.

Kapoor, S., & Narayanan, A. (2022). Leakage and the reproducibility crisis in ML-based science. *arXiv preprint arXiv*, *2207*, 07048.

Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. *Annual Review of Public Health*, *34*, 119–138.

Lansford, J. E., Godwin, J., Al-Hassan, S. M., Bacchini, D., Bornstein, M. H., Chang, L., Chen, B.-B., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Alampay, L. P., Uribe Tirado, L. M., & Zelli, A. (2018b). Longitudinal associations between parenting and youth adjustment in twelve cultural groups: Cultural normativeness of parenting as a moderator. *Developmental Psychology*, *54*(2), 362–377. https://doi.org/10.1037/dev0000416.

Lansford, J. E., Rothenberg, W. A., & Bornstein, M. H. (Eds.). (2021a). *Parenting across cultures across development: Parenting from childhood to adolescence in nine countries*. Routledge.

Lansford, J. E., Rothenberg, W. A., & Bornstein, M. H. (2021c). Change over time in four domains of parenting in diverse international contexts: Theoretical and methodological introduction. In J. E. Lansford, W. A. Rothenberg, & M. H. Bornstein (Eds.), *Parenting across cultures across development: Parenting from childhood to adolescence in nine countries* (pp. 1–21). Routledge.

Lansford, J. E., Rothenberg, W. A., Jensen, T. M., Lippold, M. A., Bacchini, D., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Uribe Tirado, L. M., Alampay, L. P., & Al-Hassan, S. M. (2018a). Bidirectional relations between parenting and behavior problems from age 8 to 13 in nine countries. *Journal of Research on Adolescence*, *28*(3), 571–590. https://doi.org/10.1111/jora.12381.

Lansford, J. E., Rothenberg, W. A., Riley, J., Uribe Tirado, L. M., Yotanyamaneewong, S., Alampay, L. P., Al-Hassan, S. M., Bacchini, D., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Gurdal, S., Liu, Q., Long, Q., Malone, P. S., Oburu, P., Pastorelli, C., & Steinberg, L. (2021b). Longitudinal trajectories of four domains of parenting in relation to adolescent age and puberty in nine countries. *Child Development*, *92*(4), e493–e512. https://doi.org/10.1111/cdev.13526.

Lavigne, J. V., Meyers, K. M., & Feldman, M. (2016). Systematic review: Classification accuracy of behavioral screening measures for use in integrated primary care settings. *Journal of Pediatric Psychology*, *41*(10), 1091–1109. https://doi.org/10.1093/jpepsy/jsw049.

Lebrun-Harris, L. A., Ghandour, R. M., Kogan, M. D., & Warren, M. D. (2022). Five-year trends in US children’s health and well-being, 2016–2020. *JAMA Pediatrics*, *176*(7), e220056 https://doi.org/10.1001/jamapediatrics.2022.0056.

Leventhal, T., & Dupéré, V. (2019). Neighborhood effects on children’s development in experimental and nonexperimental research. *Annual Review of Developmental Psychology*, *1*, 149–176. https://doi.org/10.1146/annurev-devpsych-121318-085221.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*.

Pastorelli, C., Barbaranelli, C., Cermak, I., Rozsa, S., & Caprara, G. V. (1997). Measuring emotional instability, prosocial behavior and aggression in pre-adolescents: A cross-national study. *Personality and Individual Differences*, *23*(4), 691–703. https://doi.org/10.1016/S0191-8869(97)00056-1.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*(85), 2825–2830.

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133. https://doi.org/10.1007/BF01537962.

Pettit, G. S., Harris, A. W., Bates, J. E., & Dodge, K. A. (1991). Family interaction, social cognition, and children’s subsequent relations with peers at kindergarten. *Journal of Social and Personal Relationships*, *8*, 383–402.

Phillips, L. H., Wynn, V. E., McPherson, S., & Gilhooly, K. J. (2001). Mental planning and the Tower of London task. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *54A*(2), 579–597. https://doi.org/10.1080/02724980042000237.

Pinquart, M. (2017a). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. *Developmental Psychology*, *53*(5), 873–932. https://doi.org/10.1037/dev0000295.supp. (Supplemental).

Pinquart, M. (2017b). Associations of parenting dimensions and styles with internalizing symptoms in children and adolescents: A meta-analysis. *Marriage & Family Review*, *53*(7), 613–640. https://doi.org/10.1080/01494929.2016.1247761.

Racine, N., McArthur, B. A., Cooke, J. E., Eirich, R., Zhu, J., & Madigan, S. (2021). Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19. *JAMA Pediatrics*, *175*(11), 1142 https://doi.org/10.1001/jamapediatrics.2021.2482.

Rohner, R. P. (2005). *Parental Acceptance-Rejection Questionnaire (PARQ): Test Manual*. University of Connecticut.

Rothenberg, W. A. (2019). A review of intergenerational continuity in parenting: Identifying developmental pathways and moderating factors. *Marriage & Family Review*, *55*(8), 701–736. https://doi.org/10.1080/01494929.2019.1589618.

Rothenberg, W. A., Lansford, J. E., Bacchini, D., Bornstein, M. H., Chang, L., Deater, D. K., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Tirado, L. M. U., Yotanyamaneewong, S., Alampay, L. P., & Al, H. S. M. (2020a). Cross-cultural effects of parent warmth and control on aggression and rule-breaking from ages 8 to 13. *Aggressive Behavior*, *46*(4), 327–340. https://doi.org/10.1002/ab.21892.

Rothenberg, W. A., Lansford, J. E., Al-Hassan, S. M., Bacchini, D., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Maria Uribe Tirado, L., Yotanyamaneewong, S., & Peña Alampay, L. (2020b). Examining effects of parent warmth and control on internalizing behavior clusters from age 8 to 12 in 12 cultural groups in nine countries. *Journal of Child Psychology and Psychiatry*, *61*(4), 436–446. https://doi.org/10.1111/jcpp.13138.

Rothenberg, W.A., Lansford, J.E., Uribe Tirado, M.L., Yotanyamaneewong, S., Peña Alampay, L., Al-Hassan, S.M., Bacchini, D., Chang, L, Deater-Deckard, K., Di Giunta, L., Dodge, K.A., Gurdal, S., Liu, Q., Long, Q., Oburu, P., Pastorelli, C., Skinner, A.T., Sorbring, E., Tapanya, S., Steinberg, L., & Bornstein, M.H. (2023). The intergenerational transmission of maladaptive parenting and its impact on child mental health: Examining cross-cultural mediating pathways and moderating protective factors. *Child Psychiatry and Human Development*. Advanced online publication. https://doi.org/10.1007/s10578-021-01311-6.

Singelis, T. M., Triandis, H. C., Bhawuk, D. P. S., & Gelfand, M. (1995). Horizontal and vertical dimensions of individualism and collectivism: A theoretical and measurement refinement. *Cross-Cultural Research*, *29*, 240–275.

Skinner, A. T., Bacchini, D., Lansford, J. E., Godwin, J., Sorbring, E., Tapanya, S., Uribe Tirado, L. M., Zelli, A., Alampay, L. P., Al-Hassan, S. M., Bombi, A. S., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., Miranda, M. C., Oburu, P., & Pastorelli, C. (2014). Neighborhood danger, parental monitoring, harsh parenting, and child aggression in nine countries. *Societies*, *4*(1), 45–67.

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar De Pablo, G., Il Shin, J., Kirkbride, J. B., Jones, P., Kim, J. H., Kim, J. Y., Carvalho, A. F., Seeman, M. V., Correll, C. U., & Fusar-Poli, P. (2021). Age at onset of mental disorders worldwide: Large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*, *27*(1), 281–295. https://doi.org/10.1038/s41380-021-01161-7.

Steinberg, L., Dornbusch, S. M., & Brown, B. B. (1992). Ethnic differences in adolescent achievement: An ecological perspective. *American Psychologist*, *47*(6), 723–729. https://doi.org/10.1037/0003-066X.47.6.723.

Strauman, T. J. (2017). Self-regulation and psychopathology: Toward an integrative translational research paradigm. *Annual Review of Clinical Psychology*, *13*, 497–523. https://doi.org/10.1146/annurev-clinpsy-032816-045012.

Sun, X., Ram, N., & McHale, S. M. (2020). Adolescent family experiences predict young adult educational attainment: A data-based cross-study synthesis with machine learning. *Journal of Child and Family Studies*, *29*(10), 2770–2785. https://doi.org/10.1007/s10826-020-01775-5.

Tampu, I. E., Eklund, A., & Haj-Hosseini, N. (2022). Inflation of test accuracy due to data leakage in deep learning-based classification of OCT images. *arXiv preprint arXiv*, *2202*, 12267.

Tate, A. E., McCabe, R. C., Larsson, H., Lundström, S., Lichtenstein, P., & Kuja-Halkola, R. (2020). Predicting mental health problems in adolescence using machine learning techniques. *PloS One*, *15*(4), e0230389.

Triandis, H. C. (1995). *Individualism and collectivism*. Westview.

Vallat, R. (2018). Pingouin: Statistics in Python. *The Journal of Open Source Software*, *3*(31), 1026.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. APA PsycTests. https://doi.org/10.1037/t15170-000.

Zhang, N., Liu, C., Chen, Z., An, L., Ren, D., Yuan, F., & He, G. (2019). Prediction of adolescent subjective well-being: A machine learning approach. *General Psychiatry*, *32*(5), e100096.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

W. Andrew Rothenberg is a Research Scientist affiliated with the Center for Child and Family Policy at Duke University in Durham, NC, USA and the University of Miami Miller School of Medicine's Mailman Center for Child Development, USA. His major interests include the intergenerational transmission of parenting and family processes in cultures around the world.

Andrea Bizzego is a Postdoctoral Researchers at the University of Trento, Italy. His research interests include applying machine learning algorithms and artificial intelligence to understand human development.

Gianluca Esposito is a Professor of Child Development at the University of Trento, Italy. His research interests include clinical neuroscience, comparative psychology, and developmental psychology.

Jennifer E. Lansford is Research Professor in the Sanford School of Public Policy and Director of the Center for Child and Family Policy at Duke University in Durham, NC, USA. Her research focuses on parenting and child development in diverse cultural contexts.

Suha M. Al-Hassan is an Associate Professor of Early Childhood Special Education and Applied Behavior Analysis at Emirates College for Advanced Education, United Arab Emirates, and the Hashemite University, Jordan. Her research focus includes parenting and child development, school readiness, behavior management, inclusion, and cross cultural research.

Dario Bacchini is a Full Professor of Developmental Psychology and Coordinator of the PhD Course in Mind, Gender, and Language at the University of Naples Federico II, Italy. His research focuses on antisocial behavior in adolescence, moral development, school bullying, and parenting.

Marc H. Bornstein is a senior advisor for research for ECDI Parenting Programmes at UNICEF, an affiliate with the Eunice Kennedy Shriver National Institute of Child Health and Human Development, and International Research Fellow at the Institute for Fiscal Studies in London, England. His research focuses on the experimental, methodological, comparative, developmental, and cultural science of child development.

Lei Chang is Chair Professor of Psychology and Head of Department of Psychology, University of Macau. He conducts research in the areas of evolutionary psychology, including cultural evolution, life history, and mating research, as well as developmental psychology focusing on parenting and child and adolescent social development.

Kirby Deater-Deckard is Professor in Psychological and Brain Sciences at the University of Massachusetts Amherst, MA, USA. His research focuses on the transactions between biological, cognitive, and environmental factors in the development and intergenerational transmission of self-regulation and related psychological and health outcomes.

Laura Di Giunta is Associate Professor at Sapienza University of Rome, Italy. Her research focuses on the continuity and change of individual differences in predicting youth (mal)adjustment, accounting for socialization and culture factors, especially in emotion regulation, adjustment, and social competence.

Kenneth A. Dodge is the Pritzker Professor of Public Policy and Professor of Psychology and Neuroscience at Duke University. He studies the development and prevention of aggressive and violent behaviors. His work provides a model for understanding how some young children grow up to engage in aggression and violence and provides a framework for intervening early to prevent the costly consequences of violence for children and their communities.

Sevtap Gurdal is a Senior Lecturer at the Centre for Child and Youth Studies at University West, Sweden. Her research interests include parenting and child development.

Qin Liu is an Associate Professor at Chongqing Medical University. Her research interests include the application of quantitative and qualitative methods in evaluating health services.

Qian Long is an Associate Professor of Global Health at Duke Kunshan University. Her research interests include investigating health equity in relation to healthy systems development with a focus on health financing and health services organization and delivery.

Paul Oburu is an Associate Professor in Child Development at Maseno University, Kenya. His research interests include parenting, caregiving stress, mental health, and adjustment problems of grandmother caregivers and orphaned children.

Concetta Pastorelli is a full professor at the Department of Psychology of Sapienza University of Rome, and Director of the Interuniversity Centre for the Research in the Genesis and Development of Prosocial and Antisocial Motivations (CIRMPA). Her research focuses on the determinants of aggressive and prosocial behaviors and school achievement, parental personality and parenting, and children's and adolescents' adjustment.

Ann T. Skinner is a Research Scientist at the Duke University Center for Child and Family Policy in Durham, NC, USA. Her research focuses on parenting stress and child adjustment, and the development of aggression in youth.

Emma Sorbring is a Professor of Child and Youth Studies and research director for the Centre for Child and Youth Studies at University West, Sweden. Her research and teaching interests focus on children, adolescents, and families, including young people's decision-making, teenagers' internet use and parental strategies.

Sombat Tapanya is a clinical psychologist and an independent researcher at the Peace Culture Foundation, Chiang Mai, Thailand. His research and clinical practice focuses on psychological trauma resulting from harsh parenting and abuse, and the treatment and prevention of such trauma.

Laurence Steinberg is a Professor of Psychology at Temple University and an affiliate of King Abdulaziz University, Saudi Arabia. His research focuses on psychological and brain development in adolescence.

Liliana Maria Uribe Tirado is a Research Professor at Universidad de San Buenaventura, Medellin, Colombia. She is the principal investigator of the Colombian unit in the Parenting Across Cultures Study. She is actively involved in testing evidence-based prevention programs in different regions of Colombia. Her main research focuses on parenting and child adjustment in different cultural contexts.

Saengduean Yotanyamaneewong is a clinical psychologist and a lecturer in the Psychology Department, Faculty of Humanities, Chiang Mai University, Thailand. Her research and teaching interests focus on parenting, parental burnout, and emotional and social development in children and adolescents.

Liane Peña Alampay is a Professor of Psychology at the Ateneo de Manila University, Philippines. Her research focuses on parenting and child development in cultural contexts, and interventions for youth and families at risk.