Doctoral theses at NTNU, 2024:14

Ida Sund Morken

Depression from preschool to adolescence – stability, personality, and gender

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Social and Educational Sciences
Department of Psychology

NTNU
Norwegian University of
Science and Technology

Ida Sund Morken

Depression from preschool to adolescence – stability, personality, and gender

Thesis for the Degree of Philosophiae Doctor

Trondheim, January 2024

Norwegian University of Science and Technology
Faculty of Social and Educational Sciences
Department of Psychology

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

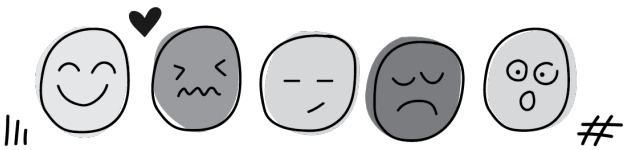
Thesis for the Degree of Philosophiae Doctor

Faculty of Social and Educational Sciences
Department of Psychology

ISBN 978-82-326-7630-9 (printed ver.)
ISBN 978-82-326-7629-3 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2024:14

Printed by NTNU Grafisk senter



NTNU

Tidlig trygg i Trondheim

Abstract

In the transition from childhood to adolescence, there is a stark increase in depression, and a female preponderance emerges. Evidently, preventative efforts are of great importance. To that end, knowledge of potential risk and vulnerability factors is needed. This thesis focused on the stability of depression from preschool into adolescence, the relation between adolescent depression and personality traits, and the potential role of stressful life events (SLEs) and bullying victimization on the female preponderance in depression. Existing research on these matters has substantial limitations.

The current work is based on cohort data from the Trondheim Early Secure Study. DSM-5-defined major depressive disorder (MDD) and dysthymia were measured biannually with clinical interviews. Predictions were investigated by using within-person methodology, which accounts for time-invariant confounding effects. In Study I (n = 1,042), we investigated five types of stability in depression from preschool into adolescence (ages 4 to 14). In Study II (n = 817), we examined whether changes in the Big Five personality traits predicted increases in depressive symptoms in adolescence—and vice versa—across ages 10 to 16. In Study III (n = 748, ages 8 to 14), we investigated if girls become exposed to more SLEs and bullying victimization than boys, and if so: whether a gendered stress exposure mediated the gender difference in depression. We also investigated whether predictions from these stressors on depression were stronger for girls than boys in early adolescence.

In Study I, we found that childhood depression increased the risk of adolescent depression, more strongly when compared to ones' own level than with others' depression level. Increases or decreases in depression predicted later corresponding changes, in line with the proposal that depression might lead to changed characteristics in the child or adolescent and/or their environment which thereafter increases the risk of recurrence (the scar hypothesis). We also found evidence for structural stability, indicating that the DSM-5 criteria

for MDD and dysthymia represent the same underlying constructs across childhood and adolescence. The most frequent symptoms were irritability and concentration difficulties. Notably, all types of stability increased when entering adolescence.

In Study II, we found that increased neuroticism predicted increases in depressive symptoms, which is in line with a predisposition explanation for the relationship between depression and personality. Moreover, increases in depressive symptoms predicted increased neuroticism and decreased extraversion and conscientiousness—in line with the scar explanation. Our findings suggest that these explanatory models may be in play already from the age of 10.

Beyond replicating the female preponderance in early adolescent depression, Study III showed that increased SLEs and bullying victimization predicted increased depression more strongly for girls than boys—in early adolescence (i.e., just prior to the emerging gender difference)—but not late childhood. Girls did not become exposed to more SLEs or to more bullying victimization than boys. These results indicate support for the stress reactivity—but not the stress exposure—explanation for the female preponderance in adolescent depression.

Collectively, our findings underscore the transition from childhood into early adolescence as a period of risk for depression. All types of stability increase at this age, and depression might have an impact on several personality traits. Our findings also suggest that early adolescents with a history of early-onset depression as well as those scoring high on neuroticism might me especially vulnerable to developing depression, and early adolescent girls may react more strongly with increased levels of depression when exposed to SLEs and bullying victimization. Childhood depression, high levels of neuroticism, and gendered stress-reactivity may be accounted for in treatment and preventative efforts.

Contents

Acknowledgements

The heterogenous, draining, and passive nature of depression has, paradoxically so, long fascinated me. I am very grateful for the opportunity to learn more about and contribute with a piece of the depression-puzzle. There are several groups and people to be acknowledged:

The Department of Psychology, NTNU

All work presented in the current thesis were facilitated and funded by the Norwegian University and Technology, Department of Psychology. The Trondheim Early Secure Study (TESS) is funded by grants from the Research Council of Norway and the Liaison Committee between Central Norway Regional Health Authority and NTNU.

The TESS

Thank you to all the children, parents, and teachers participating in the TESS for your openness and valuable time, wave after wave.

I would also like to thank all my colleagues in TESS research group. A special thanks to the research assistants for their overview, thoroughness, and meticulous performance in interviewing children and their parents. Thank you Tilmann von Soest, Silje Steinsbekk, and Bror Morten Ranum for very important contributions to my papers.

Supervisors

To my supervisors Kristine Rensvik Viddal and Lars Wichstrøm. Your profound curiosity in development of depression has strongly contributed to an enjoyable PhD-process and is something I hope to bring with me as a researcher.

Lars Wichstrøm: I feel so lucky to have been able to draw on your substantial experience in investigating depression in children and adolescents, as well as statistical expertise. Thank you for your calm nature, patience, and all the knowledge you have passed on.

To Kristine Rensvik Viddal. Thank you for your presence, pragmatism, and sharing from your vast archive of theoretical knowledge. And for our long discussions, first and foremost about our work, but also about life's mysteries and odd practicalities. Thank you for being an ever-safe haven.

Other important people and locations

- My friends. Forever grateful.
- To my sisters Tora and Mari, and Anders & Anders—thank you for your support and dinners. To my various nieces and nephews spanning even a wider age-range than the TESS: thank you for your vitality, presence, and insight.
- Thank you to my parents, who have always firmly believed in my abilities. By providing me with a variety of research assistant jobs from age 15, you have made doing research seem not that scary, and even attainable.
- To my colleagues at the Journal of the Norwegian Psychological Association, for all your interest and encouragement.
- Sagveien 23A, Oslo Tango, Tranen, Rouleur, Libros del Pasaje, Bar de Fondo, Cariño.
- Evalill, Hanne, Espen, Sturla, Ingalill, Leif, and Steffen.

Oslo, September, 2023

Ida Sund Morken

List of studies

Morken, I. S., Viddal, K. R., Ranum, B., & Wichstrøm, L. (2020). Depression from preschool to adolescence—five faces of stability. *Journal of Child Psychology and Psychiatry*. *62*(8), 1000-1009. https://doi.org/10.1111/jcpp.13362

Morken, I. S., Wichstrøm, L., Steinsbekk, S., & Viddal, K. R. Depression and Personality Across Adolescence—Within-person Analyses of a Birth Cohort. *Submitted.*

Morken, I. S., Viddal, K. R., Von Soest, T., & Wichstrøm, L. (2023). Explaining the Female Preponderance in Adolescent Depression—A Four-Wave Cohort Study. *Research on Child and Adolescent Psychopathology*, 1-11. https://doi.org/10.1007/s10802-023-01031-6

Acronyms and abbreviations

ADHD	Attention-Deficit/Hyperactivity Disorder
AIC	Akaike Information Criterion
ALT-SR	Autoregressive Latent Trajectory Models with Structural Residuals
APA	American Psychological Association
BFI	Big Five Inventory
CAPA	Child and Adolescent Psychiatric Assessment
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CI	Confidence Interval

CLPM	Cross-Lagged Panel Model
DAWBA	Development and Well-Being Assessment
df	Degrees of Freedom
DHEA	Dehydroepiandrosterone
DSM	Diagnostic and Statistical Manual of Mental Disorders
DMDD	Disruptive Mood Dysregulation Disorder
FFM	Five Factor Model
FIML	Full Information Maximum Likelihood
GH	Growth Hormone
HPA-axis	Hypothalamus-Pituitary-Adrenal axis
HPG-axis	Hypothalamic-Pituitary-Gonadal axis
ICC	Intra-Class Correlation
ICD	International statistical Classification of Diseases
k	Kappa
Kiddie-SADS	Schedule for Affective Disorders and Schizophrenia for School-Age Children
MAR	Missing At Random
M_{age}	Mean age
MDD	Major Depressive Disorder
NDPA	Norwegian Data Protection Authority
OBVQ	Olweus Bully Victim Questionnaire
ODD	Oppositional Defiant Disorder
OR	Odds Ratio

PAPA	Preschool Age Psychiatric Assessment
PDD	Persistent Depressive Disorder
r	Pearson product moment correlation coefficient
r^2	Pearson *r* squared
REK	Regional committee for medical and health research ethics
RI-CLPM	Random Intercept Cross-Lagged Panel Model
RMSEA	The Root Mean Square Error of Approximation
SD	Standard Deviation
SDQ	Strengths and Difficulties Questionnaire
SE	Standard Error
SES	Socioeconomic Status
SLEs	Stressful Life Events
TESS	Trondheim Early Secure Study
$\chi 2$	Chi square
$\Delta\chi 2$	Delta chi square

1. INTRODUCTION

1.1. Current context and focus

Depression is a serious health issue affecting over 300 million people worldwide (World Health Organization, 2017) and is one of the leading causes of years lived with disabilities (Global Burden of Disease, 2022). First onset typically occurs in early adolescence (e.g., Merikangas et al., 2010a). Moreover, in this age period, a female preponderance in depression emerges (Salk et al., 2017). Adolescent depression confers a risk for myriad long-term detrimental outcomes, such as failure to complete secondary school, unemployment, and poor physical health and social functioning in adulthood (Clayborne et al., 2019; Copeland et al., 2021). Even subclinical levels of depression may lead to future negative outcomes, such as major depressive disorder (MDD), as well as lower quality of life (Bertha & Balázs, 2013) and earnings (Hakulinen et al., 2016).

During the last decade, the prevalence rates of depressive disorders and symptoms have likely increased (Moreno-Agostino et al., 2021), including among adolescents (for a meta-analysis, see Shorey et al., 2022). The same increase has been shown among adolescents in Norway (Bakken, 2020; Bråten et al., 2023)—where the present work was conducted. Moreover, studies indicate that the gender difference in adolescent depression might be widening (Bakken, 2020; Bråten et al., 2023; Platt et al., 2021). As such, preventing adolescent depression—including the female preponderance—is crucial. To efficiently plan and execute both preventative and treatment interventions, we must understand the etiology of adolescent depression. Despite the substantial research literature on depression, crucial knowledge gaps exist—especially from a developmental point of view.

Importantly, for many individuals, depression is a recurrent condition (Burcusa & Iacono, 2007; Kovacs et al., 2016). Indeed, earlier depression increases the risk of recurrence both during adulthood (Richards, 2011) and from adolescence into adulthood (Copeland et al.,

2021; Johnson et al., 2018). Depressive disorders may also be identified in the preschool period (Vasileva et al., 2021). However, whether they reappear in adolescence has not been thoroughly investigated. Moreover, we do not know whether the depression construct measured in adolescents is the same when measured in children. To elucidate these questions, different forms of stability in depression were investigated in **Study I** across ages 4 to 14.

As noted above, uncovering the factors that might contribute to the development of adolescent depression is critical. In observational research, it is necessary that these factors (e.g., childhood emotional neglect, Glickman et al., 2021; premature birth ElHassan et al., 2023) present as *statistical* risk factors. Characteristic ways of feeling, acting, and thinking—often defined as *personality* (Allport, 1961)—could confer one such factor. Indeed, particularly high or low levels of some personality traits as captured in the Five Factor Model (FFM; John, et al., 2008) have been suggested to play an etiological role in adolescent depression (Klimstra et al., 2010). This causal direction is captured in the *predisposition model* (Ormel et al., 2013). Furthermore, adolescent depression might impact personality development. The *scar model* entails that depression leaves changes in personality traits after the depressive episode has subsided (Ormel et al., 2013). Whether the predisposition and scar models apply in adolescence—and already from the earliest adolescent years—is undetermined. In **Study II,** we investigated whether changes in the Big Five personality traits predicted increases in depressive symptoms—as well as the reverse—across ages 10 to 16.

Finally, several explanatory models have been developed to account for the female preponderance in depression (e.g., Cyranowski et al., 2000; Hankin & Abramson, 2001; Hyde & Mezulis, 2020). The *stress exposure model* proposes that when entering adolescence, girls become exposed to more stress than boys—and therefore more depressed (Hammen, 2009b; Hankin et al., 2007). A related explanation is the *stress reactivity model*: girls react more strongly to stress than boys do, in the form of more depression (Hammen, 2009b; Hankin et

al., 2007). Two well-established risk factors for adolescent depression are stressful life events (SLEs; Ge et al., 1994) and bullying victimization (Christina et al., 2021). However, whether increased exposure or reactivity to these stressors contribute to explaining the gender difference in depression is uncertain. In **Study III,** we examined if girls become exposed to more stress than boys in early adolescence and whether gendered stress exposure mediates the gender difference in depression. We also investigated whether increased stress predicted increases in depressive symptoms to a stronger degree in girls than boys—and more so in early adolescence than late childhood.

Despite theoretical and research efforts to gain knowledge about the development of adolescent depression—including the gender difference—many unanswered questions remain. When compared to the research base on adult depression, considerably fewer studies focus on childhood and adolescence—and existing studies have important limitations. First, previous longitudinal studies have almost exclusively applied traditional regression-based methods. These examine whether an individual's level of a risk (or protective) factor—compared to others' levels of this factor—predicts an individual's depression level—compared to others' depression levels. Thus, these methods render it unclear whether longitudinal associations reflect between-person differences in levels of the predictor and depression, or within-person changes in the predictor and depression (Hamaker, et al., 2020). Moreover, results from these studies do not necessarily imply that a predictor explains (i.e., causes) adolescent depression; rather, it may appear due to confounding effects stemming from unmeasured factors (Hamaker et al., 2020). Confounding effects may be time variant, stemming, for example, from a newly occurring life event (Ge et al., 1994). They may also be time invariant, stemming, for example, from relatively stable parental practices increasing the likelihood of both a predictor (e.g., bullying victimization) and later depression to occur (Tang et al., 2018). By contrast, within-person methodology disentangles between- and

within-person information by using participants as their own controls and thus accounts for one of these types of confounding effects: the time-invariant (Berry & Willoughby, 2017). To date, few studies on child and adolescent depression have applied such methods—and none directly inform the research questions in the current thesis. In our studies, we applied within-person methodology and therefore informed on possible explanations for adolescent depression to a *stronger* degree than previous studies. Evidently, though, time-variant confounding persists. Notwithstanding, given that observational methods are not able to meet all assumptions for a causal relationship, causal inferences are not possible (Mund & Nestler, 2019). As such, the results from the studies included in this thesis inform on the *likelihood* of etiology.

A second limitation of the prior longitudinal studies is that they almost exclusively began in early adolescence. The typical baseline age is 12–13 years—paralleling the age that depression first becomes more prevalent (Merikangas et al., 2010a) and the female preponderance first emerges (Salk et al., 2017). However, changes in a predictor—such as a personality trait or a stressor—should occur *before* the increase in the outcome (depression). Personality traits may temporarily change while an adolescent is depressed (Ormel et al., 2013), and depression may generate stress—at least partly dependent on the adolescent's characteristics and behavior (Hammen, 2009a). Hence, changes in these factors from ages 12–13 and onward do not imply that the personality trait or stressor are unequivocal predictors of increased depression. The changes in these factors could rather be effects of increased depression during the same period. Thus, to explain adolescent depression—including the gender difference—occurring as young as ages 12–13, it is important to design studies with a younger baseline age.

1.1.1. Overarching research questions

Based on the empirical and methodological issues presented above, this thesis aims to answer the following questions:

a) Is depression a stable phenomenon across preschool into adolescence?

b) Do personality traits predict changes in adolescent depression? Furthermore, does adolescent depression predict changes in personality traits?

c) Do girls become exposed to more stress in the transition to adolescence? If so: does such a gendered stress exposure mediate the gender difference in depression? And/or does stress predict depression to a stronger degree for girls than boys?

1.2. Childhood and adolescent depression

Depression is characterized by persistent feelings of sadness, irritability, and/or loss of interest in things and activities the child or adolescent used to enjoy. It may leave the child or adolescent more socially withdrawn, with feelings of worthlessness and guilt, suicidal thoughts, and, in some cases, even suicide attempts (e.g., Bernaras et al., 2019). Historically, depression has been conceptualized and understood from differing perspectives, such as psychoanalysis, behaviorism, interpersonal theories, cognitive psychology—and biological theories (for an overview, see Carr, 2015). In current-day research—and clinical work—diagnostic systems are often used to identify and classify depression. The Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013) and the International Classification of Diseases (ICD-11; World Health Organization, 2022) cover mental disorders. These systems have been developed according to international expert working groups named by the American Psychiatric Association and the World Health Organization, respectively. In the studies included in the current thesis, we measured depression according to the DSM-5, which is presented in Section 1.2.1.1.

Importantly, these diagnostic systems are purely descriptive, and a depression diagnosis may entail different symptom constellations for different individuals. Accordingly, depression is considered a syndrome (a co-occurring collection of symptoms; Kendler & First, 2010) and not a disease (a condition with an identified cause in the organism; Miriam-Webster, 2023). Therefore, the diagnostic systems cannot precisely define how depression will present for a given individual, nor can they inform on the underlying causes or sustaining mechanisms.

1.2.1. Conceptualization and measurement

1.2.1.1. The construct of depression according to the DSM-5

According to the DSM-5, depressive disorders have one common feature: a sad or irritable mood and/*or* anhedonia. In the current work, depression was operationalized as major depressive disorder (MDD) and dysthymia—as defined by the DSM-IV—in the first waves of the study beginning in 2007. In 2013, DSM-5 was published. Dysthymia, chronic minor MDD, and recurrent MDD without recovery between episodes now fall under the term persistent depressive disorder (PDD; Schramm et al., 2020). An important difference from DSM-IV-defined dysthymia is that during the PDD episode, MDD may be continuously present. Even so, the dysthymic symptoms remained unchanged in the DSM-5 (American Psychiatric Association, 2013). The studies included in the current thesis include measures of depressive disorders before (ages 4, 6, 8 and 10) and after (ages 12, 14 and 16) the introduction of DSM-5. Given that the symptoms described in DSM-IV and -5 are the same, we consequently use the term “dysthymia” in the studies.

MDD and dysthymia are the two most common depressive disorders, with some overlap in symptoms (see Table 1). In MDD, the core symptoms (i.e., symptoms that must be present for a diagnosis) are depressed mood or loss of pleasure/interest (anhedonia). In dysthymia, the core symptom is depressed mood. Which and how many of the other

symptoms—as well as how long they last—determines whether the episode is classified as MDD or dysthymia.[1]

Table 1: DSM-5 symptoms of MDD and dysthymia

Symptoms	MDD	Dysthymia
Depressed mood/irritability	✓	✓
Significant weight loss (appetite) or gain	✓	✓
Insomnia or hypersomnia	✓	✓
Fatigue or loss of energy	✓	✓
Diminished ability to think or concentrate, or indecisiveness	✓	✓
Anhedonia/marked diminished interest or pleasure in most or all activities	✓	
Psychomotor agitation/retardation	✓	
Feelings of worthlessness or excessive or inappropriate guilt	✓	
Recurrent thoughts of death (not just fear of dying) or suicidal ideation, plan, or attempt	✓	
Low self-esteem		✓
Feelings of hopelessness		✓

Source: DSM-5; American Psychiatric Association (2013).

When applied to children and adolescents, there are some adjustments to the DSM-5 criteria for adults. First, the core symptoms are extended from depressed mood (for MDD and dysthymia) or loss of interest/pleasure (for MDD) in adults to also include persistent irritability. Second, significant weight loss or gain may be replaced by "failure to meet expected weight gains". Third, for dysthymia, the duration of symptoms must be at least one

[1] The symptoms should not be attributed to the physiological effects of a substance or a medical condition, nor to a psychotic disorder (American Psychiatric Association, 2013).

year (two years in adults). Of note, in the interview we used in Study I to measure MDD at ages 4 and 6, as proposed by Luby et al. (2002) age-appropriate criteria for suicidality were added: persistent preoccupation with death or self-harm in play. Luby et al. (2002) found empirical support for these criteria, for example, in that children diagnosed with depression had higher levels of death-related or suicidal themes in play (61%) when compared to normal and externalizing disorder control groups. For the sake of order, no such adjustments for children or adolescents are made in the ICD-11. Thus, the ICD-11 implicitly conveys that depression appears the same in children, adolescents, and adults.

1.2.1.2. Categorical and dimensional conceptualizations

When applying a dimensional conceptualization—which is common within the field of developmental psychopathology (Cicchetti & Toth, 2009)—depressive symptoms are not categorically separated into clinical versus subclinical depression. The underlying measurement assumption is that depressive symptoms are quantitative, continuous, and linear (Essau & Ollendick, 2009). Thus, a dimensional approach differentiates between individuals with various degrees or numbers of depressive symptoms, including those that do not reach a clinical threshold. More cases are included, which increases the variability in depression scores and thus statistical power (Field 2013). Increased power reduces the risk for Type 2 errors, or false negatives (Field, 2013). Given that prior research has shown a very low prevalence of depression in the preschool period (Vasileva et al., 2021), this advantage is especially relevant for Study I in which we investigated depression from age 4.

Finally, research has shown that even subclinical levels of depression are associated with impaired familial, school, and peer functioning, other psychiatric disorders, and suicide (Gibb, 2014; Rudolph & Flynn, 2014). Thus, preventing depressive symptoms before they reach a level at which they warrant a clinical diagnosis is important. Studies using a

dimensional approach inform on the development of subthreshold depressive symptoms and may therefore provide knowledge relevant for preventative efforts.

1.2.1.3. Questionnaires versus clinical interviews in research

In research on categorically defined depression, clinical interviews are often used, which include specific diagnostic criteria. Interviews are costly and time-consuming to administer, but they also have several advantages. For example, they allow the interviewer to register symptoms that the person afflicted by depression may not be aware of (Essau & Ollendick, 2009). An example is psychomotor retardation, i.e., the slowing down of body movements. Moreover, when interviews are structured or semi-structured, there is less risk that different interviewers will ascribe different degrees of clinical significance to a particular symptom (Le Couteur & Gardner, 2008). Finally, because the semi-structured format in particular allows for follow-up questions, resolves inconsistencies, and elicits examples (Klein, 2023). They provide more information about the quality and intensity of the depressive symptoms than is possible with structured interviews and questionnaires.

By contrast, studies with a dimensional conceptualization of depression often use questionnaires (Essau & Ollendick, 2009; Verhulst & Van der Ende, 2008). Questionnaires have several advantages: They are inexpensive and easy to administer; they often have normative data to determine severity (Essau & Ollendick, 2009; Verhulst & Van der Ende, 2008); and most are psychometrically sound as regards reliability and validity (Stockings et al., 2015). However, there are several limitations. Many questionnaires include items that do *not* describe depressive symptoms as defined by the DSM (Bernaras et al., 2019), e.g., "I didn't have any fun at school" (Costello & Angold, 1988). They also lack information on the onset, duration, and frequency of depressive symptoms (Essau & Ollendick, 2009). Stockings et al. (2015) conducted a meta-analysis including the most common questionnaires and found evidence for moderate sensitivity (proportion of depressed children correctly assigned with a

diagnosis) and specificity (correctly *not* assigned) (e.g., Shean & Baldwin, 2008). Positive predictive power (the proportion of the children assigned with depression actually depressed) (Thelle & Laake, 2008) was mostly poor (Stockings et al., 2015). Finally, questionnaires are limited by the informant's perspective. Even though this may be partially amended by using multi-informant data (Verhulst & Van der Ende, 2008), this is rarely done in research on childhood or adolescent depression.

1.2.1.4. Conceptualization and measurement in the current work

As described above, there are apparent advantages to using clinical interviews in research. Moreover, a dimensional classification of depression confers increased statistical power. Given these advantages, we combined clinical interviews with a dimensional conceptualization of MDD and dysthymia. One exception was Study I (different types of stability): we applied a categorical approach when suitable to the research question (e.g., when investigating the prevalence of these disorders).

Given that our research questions pertained to the depression-phenomenon and not MDD or dysthymia specifically, we used a combined measure of symptoms of MDD and dysthymia. It follows that the implications that can be drawn concern depressive symptoms in general. Of note, though, in Study I on stability, MDD and dysthymia were investigated separately. Their stability may differ (e.g., because the duration criterion for dysthymia is at least one year). Moreover, it has been reported that one type of depression might increase the risk of another (e.g., Klein et al., 2006). Therefore, the stability of a combined measure of MDD and dysthymia may not necessarily confer on each specific depressive disorder.

Finally, in our studies, we interviewed multiple informants (parents and children), which is ideal when measuring child psychopathology (Achenbach, 2006; Achenbach et al., 1987). We were thus able to capture depressive symptoms from different vantage points while reducing the likelihood of common source bias.

1.2.2. Is depression the same construct across childhood and adolescence?

Another question relevant to the conceptualization of depression is whether one captures the same construct in different age-groups. Historically, it was thought that children could not suffer from depression (Schulterbrandt & Raskin, 1977). However, in more recent years, depression has been identified in children as young as preschool age (Lavigne et al., 2009; Vasileva et al., 2021). That said, whether depression is the same phenomenon across childhood and adolescence (Weiss & Garber, 2003)—*stability of form*—is an important but somewhat neglected issue.

Stability of form may be described two-fold: a) *structural stability*: if the symptoms that are most important to the overall depression construct are the same across developmental levels; and b) *frequency of individual symptoms*: if the symptoms that are most typical are the same across ages (Weiss & Garber, 2003). Although the frequency of individual symptoms may change with age, they could be equally important to the overall construct. Notably, the DSM-5 criteria for childhood depression (the inclusion of irritability as a core symptom, the changed weight criterium, one-year duration of dysthymia as opposed to two years for adults) are identical across earlier childhood and adolescence. Therefore, the DSM-5 presupposes structural stability across these ages. Clinically, it is important to know whether a DSM-defined symptom is in fact indicative of a depression diagnosis among youths. Also, establishing structural stability in the depression construct is a prerequisite for interpreting other types of stability. For example, if results indicate that childhood depression increases the risk for adolescent depression (i.e., rank-order stability), it is important to know whether the constructs we measure in the two developmental periods are identical.

Previous meta-analytic findings on (a) structural stability are mixed. Some studies report that symptoms are equally important to the depression construct across ages (reviewed by Weiss & Garber, 2003). Others have found guilt to be more important in children (Weiss

& Garber, 2003), and anhedonia (Weiss & Garber, 2003) and fatigue (Lahey et al., 2004; Weiss & Garber, 2003) in adolescents. Furthermore, most population-based studies have used questionnaires that include only some of the DSM-defined symptoms and, additionally: *non*-DSM symptoms. Meta-analytic results on b) frequency of symptoms indicate that some symptoms (e.g., anhedonia, social withdrawal, hypersomnia, weight gain, hopelessness) are more typical in adolescence than in earlier childhood (Weiss & Garber, 2003). However, the population-based (i.e., not clinical) studies included in this meta-analysis have used questionnaires that do not include all the DSM-defined depressive symptoms. Moreover, there is a dearth of research after the mid-1990s. Collectively, whether the structure of DSM-defined diagnoses holds across preschool into adolescence—and which symptoms are most frequent at which ages—is uncertain and was investigated in Study I.

1.2.2.1. The increased prevalence from childhood to adolescence

A fundamental question regarding the development of depression concerns how common it is at different ages. Even though depression has been identified in very young children (Lavigne et al., 2009) the prevalence is substantially higher in adolescence (Hankin et al., 2015; Merikangas et al., 2010b). However, exactly when, from preschool age to adolescence, the increase in depression occurs is somewhat unclear.

When investigating how common depression is among children (i.e., prevalence), a categorical approach is usually applied. Previous studies on the preschool period have found low prevalence rates for both MDD (0.3%) and dysthymia (0.2%-0.3%) (Lavigne et al., 2009; Wichstrøm et al., 2012). In later childhood, the prevalence is seemingly higher, with estimates for MDD between 0.3% (Ford, Goodman, & Meltzer, 2003) and 1.6% (Merikangas et al., 2010a) and dysthymia between 0.1% (Costello et al., 1996) and 0.8% (Merikangas et al., 2010a). In early adolescence, the prevalence rates for MDD vary between 3.8% (Merikangas

et al., 2010a) and 5.5% (Avenevoli et al., 2015). For dysthymia, the prevalence estimates are between 0.2% (Gau et al., 2005) and 1.1% (Merikangas et al., 2010a).

Collectively, when inspecting prevalence rates, both MDD and dysthymia seem to increase from preschool age until early adolescence. A few studies have investigated *when* across these ages, the increases occur. Only one study has examined the potential increase of MDD in earlier childhood and found no increase from ages 5-7 to 8-10 (Ford et al., 2003). In these ages, dysthymia has not been investigated. From middle childhood into adolescence, results on MDD are mixed. One study did not find any increase from ages 9 to 13 (Costello et al., 1996), while two studies indicated an increase around age 13 (an increase from ages 11-12 to 13-15; Ford et al., 2003; and from ages 8-11 to 12-15; Merikangas et al., 2010a). As regards dysthymia, two studies reported no change (Costello et al., 1996; Merikangas et al., 2010a). With the exception of one study conducted in Great Britain (Ford et al., 2003)—the studies are conducted in the United States, and research in the northern European context is lacking. Both the prevalence rate and when the potential increases in MDD and dysthymia occur were investigated as part of Study I.

Summing up, depression is more common in early adolescence than in very young children, even though the exact timing of this change in the prevalence is uncertain. Studies also show that depression increases even further into middle to late adolescence (e.g., Hankin et al., 2015; Merikangas et al., 2010b). Therefore, understanding why depression often emerges in adolescence is pertinent.

Notably, adolescence is marked by normative neurobiological, hormonal, psychological, and social changes (Cicchetti & Toth, 2009). Many of these have been suggested to increase the vulnerability for depression in the early adolescent period (see e.g., Cicchetti & Toth, 2009; Harter, 2015; Nolen-Hoeksema & Hilt, 2013; Rapee et al., 2019). Based on these prior literature reviews, I will provide examples of developmental aspects

related to puberty and psychosocial changes in the following. These developmental considerations were not investigated in the studies included in the current thesis. Yet, they provide a wider context for the development of adolescent depression and the specific research questions we did investigate.

1.3. Developmental considerations for adolescent depression

1.3.1. Pubertal development

Pubertal development involves hormonal, physical, and cognitive changes. These require a concerted effort activated by the hypothalamic-pituitary-gonadal (HPG)-axis (Bordini & Rosenfield, 2011). An overall higher estrogen level in girls and a higher testosterone level in boys facilitate the development of reproductive maturation and primary sexual characteristics (Sisk & Foster, 2004). The growth hormone (GH) drives the development of secondary sexual characteristics (Bordini & Rosenfield, 2011) and physical growth. For girls, this entails increased height and more overall fat and hip-width, and for boys: height, muscle mass, and shoulder width (Carr, 2015). Importantly, all these pubertal changes occur in the age-period in which the prevalence of depression first increases (Merikangas et al., 2010a), and the female preponderance first emerges (Salk et al., 2017). Entering puberty has been shown to increase the risk of internalizing disorders, even more so than chronological age (reviewed by Mendle, 2014 and; Pfeifer & Allen, 2021). It is therefore likely that several of the normative changes associated with puberty, including pubertal timing; changes in emotional reactivity and regulation; and sleep debt, leave adolescents vulnerable to depression.

1.3.1.1. Pubertal timing

Pubertal timing may be defined as the adolescent's relative pubertal development when compared with other same-gender, same-age, adolescents, classifying the individual as "early", "on-time", or "late" (Petersen & Taylor, 1980). In Norway, the average age for the

emergence of primary sexual characteristics in 2020 was 10.4 for girls (Bruserud et al., 2020) and 11.7 for boys (Madsen et al., 2020)—ages comparable with other northern European countries (Oehme et al., 2020). The changes in body composition also begin about one to two years earlier in girls (Carr, 2015).

Deviating pubertal timing has been proposed as a risk factor for psychopathology (Graber, 2013; Mendle & Ferrero, 2012; Reardon et al., 2009), including depression (Ge et al., 2001; Negriff & Susman, 2011; Wichstrøm, 1999). Some find early timing as a risk factor for depression only in girls (Ge et al., 2001)—while others find a risk for both genders (Negriff et al., 2008; Wichstrøm, 1999). Results from a recent meta-analysis indicate that early pubertal timing is a risk factor for the overall dimensions of internalizing and externalizing psychopathology, but *equally* so for girls and boys (Ullsperger & Nikolas, 2017). Whether there is a gender difference in risk of early timing on depression specifically should be investigated meta-analytically. As of now, available evidence indicates early timing as a risk factor for both girls and boys. However, as Ullsperger and Nikolas (2017) note, mechanisms explaining why early timing predicts depression—as well as potential moderators for this prediction—might be gender-specific. For example, gaining more weight than most same-gender peers might pose a stronger risk for depression in girls than boys (e.g., Hyde et al., 2008).

1.3.1.2. Emotional reactivity and regulation

Early adolescence is characterized by a discrepancy between *emotional reactivity* (the tendency to experience frequent and intense emotional arousal) and *response inhibition* (Rapee et al., 2019). Pubertal hormones likely influence the rapid maturation of brain areas involved in bottom-up limbic reward and socioemotional systems associated with increased emotional reactivity (Somerville et al., 2010). By contrast, the prefrontal cortex (PFC)—and associated executive cognitive functions involved in response inhibition—develop in a more

linear and gradual manner throughout adolescence and into early adulthood (Somerville et al., 2010; Steinberg, 2008). Therefore, the discrepancy between reactivity and inhibition peaks in early adolescence and decreases towards early adulthood (Rapee et al., 2019). Thus, early adolescents become rapidly emotionally activated and thereafter struggle to inhibit (i.e., down-regulate) or adjust accompanying behavioral tendencies (Arain et al., 2013). This may leave them with intense and prolonged negative affect—with obvious parallels to emotional disorders (Rapee et al., 2019; Somerville et al., 2010).

The emotion regulation (ER) strategies adolescents use may also confer vulnerability for depression. Pubertal hormones contribute to the development of brain-areas involved in higher-order cognitive processes (Larsen & Luna, 2018) even though they develop slower than the limbic system (Somerville et al., 2010). These cognitive advancements allow adolescents to more easily self-regulate than younger children (Silvers et al., 2012), including through the use of cognitive regulation strategies (Compas et al., 2017). However, such strategies may not necessarily be *adaptive* in this age. For example, a meta-analysis found that the use of cognitive reappraisal and emotional suppression predicted increased internalizing symptoms in adolescents but not in children (Compas et al., 2017). Another ER-strategy is rumination (Aldao, Nolen-Hoeksema, & Schweizer, 2010): responding to negative feelings and experiences by passively and repeatedly focusing on distress, possible causes, and consequences (Nolen-Hoeksema et al., 2008). Rumination has also been shown to predict adolescent depression (Hankin, 2008; Young & Dietrich, 2015) (for a discussion on *co*-rumination, see section 1.3.2.1.).

1.3.1.3. Sleep

In early adolescence, there is a major change in the timing of and need for sleep (e.g., Short et al., 2018). An increase in estradiol and testosterone likely delays circadian rhythms and the sleep/wake cycle (Hagenauer & Lee, 2012). Sleepiness takes longer to accumulate

during the day (Jenni et al., 2005). The changes allow adolescents to stay awake longer, with a corresponding need for a later wakening (Carskadon, 2011), while school starts at the same hour as in earlier childhood. This may explain why many adolescents have a sleep-debt (Rapee et al., 2019). Use of technological devices late at night or in bed may also lead to reduced sleep (Hysing et al., 2015). For example, both the use of a mobile phone specifically—as well as the number of devices in the bedroom—have been associated with more sleep-problems in adolescents (Bruni et al., 2015).

Sleep-debt increases the risk of adolescent depression (Short et al. 2020). The relationship between sleep-debt and depression is most likely bidirectional (Kelly & El-Sheikh, 2014)—and naturally so given that sleep disturbances are a symptom of depression (American Psychiatric Association, 2013). However, prior studies do indicate that sleep-disturbances more strongly predict adolescent depression rather than the opposite (Kelly & El-Sheikh, 2014; Lovato & Gradisar, 2014).

1.3.2. Psychosocial changes

The entrance into adolescence is marked by major psychosocial changes (Rudolph & Klein, 2009). Parents, peers, and teachers have differing expectations (Harter, 2015), while the concern with the evaluation of others increases (Harter, 2015; Westenberg et al., 2004). Thus, the social world becomes more complex. Pubertal reproductive hormones likely activate motivational tendencies towards potential romantic partners, an overall social reorientation from parents towards peers (Forbes & Dahl, 2010), as well as increased sensation-seeking: the motivational tendency to want to experience high-intensity, exiting experiences (reviewed by Bailen et al., 2019; and by Forbes & Dahl, 2010). The emotionally reactive early adolescent more often places themselves[2] in high risk-situations than younger

[2] In the papers, we used the terms "him/himself/his" or "her/herself/hers". In the thesis however, I use "they/themselves/their". This is a generic singular third-person singular pronoun, which is endorsed by the APA because it is inclusive of all people (https://apastyle.apa.org/style-grammar-guidelines/grammar/singular-they)

children (Forbes & Dahl, 2010). Previous research has focused on a range of psychosocial stressors, spanning teen pregnancy, parental divorce and sexual assault (Hammen, 2009b). The focus herein will be on stressful life events (SLEs).

SLEs are usually measured with checklists or interviews sampling a variety of events and experiences covering a broad range of seriousness and deemed especially relevant in childhood and adolescence (Grant et al., 2004). Examples of commonly included events are parental divorce, having experienced a fight with a close friend; losing a pet; moving, and having experienced violence. Most prior studies indicate an increase in SLEs from middle childhood to adolescence (Ge et al., 1994; Jenness et al., 2019; Larson & Ham, 1993), albeit with one exception (Garber et al., 2002). Furthermore, most studies report that SLEs predict depression in early adolescence (e.g., Ge et al., 2001). Whether the predictions are stronger in this age-group compared to earlier childhood has not been investigated. As of now, the evidence suggests that increased SLEs may help explain why depression becomes more prevalent in adolescence.

1.3.2.1. Peers

As noted above, the social reorientation of early adolescence entails an interest in—and pursuit of—contact with peers and potential romantic partners (Forbes & Dahl, 2010). Peers become the primary source of social support (Harter, 2015). In parallel, an advancing capacity to imagine multiple perspectives (Enright et al., 1980) likely contributes to a peak in self-consciousness (Harter, 2006) which decreases towards young adulthood (Frankenberger, 2000). These developmental changes may leave the early adolescent vulnerable to depression in different ways.

First and foremost, if the adolescent does not perceive peer social support as available, they might feel rejected (Harter, 2015). Lack of support may threaten a fundamental need to belong (Baumeister & Leary, 1995; Verhagen et al., 2018). Second, the adolescent may not

seek out support when they need it (Thompson et al., 2006). Both of these processes may increase the risk of depression (Baumeister & Leary, 1995; Thompson et al., 2006). Of note, adolescents *internal working models*—mental representations formed through early experiences with their parents (Bowlby, 1982)—might influence adolescent depression (Rudolph, 2009). Children that experience inconsequential or unreliable support, love, and attention might construe a *negative* internal working model: that they are not worth being loved, and that other people are unreliable or will not respond when they are needed (Bowlby, 1982). When these children become adolescents, they may more readily feel rejected (Rudolph & Klein, 2009) and be less likely to perceive and seek peer social support in the face of stress (Thompson et al., 2006).

Related to rejection by peers: bullying victimization is another important risk factor for adolescent depression (Christina et al., 2021). Although bullying victimization may not increase in adolescence (McLaughlin et al., 2009; Sweeting et al., 2006; Wendelborg, 2020), early adolescents display increased emotional reactivity (Rapee et al., 2019) —especially to *social* cues—when compared to children and adults (Bailen et al., 2019; Hare et al., 2008; Somerville et al., 2013). The increased importance and salience of peers (Harter, 2015; Rapee et al., 2019)—coupled with stronger emotional reactions—might leave the early adolescent more vulnerable to peer rejection and bullying victimization than younger children.

Arguably, social media introduces a widening of the peer context. Moreover, it broadens the adolescent's developmental microsystem: the relation between the developing person and their environment in the immediate setting (e.g., home, school) (Bronfenbrenner, 1977). This may have advantages (Valkenburg et al., 2022). For example, it provides opportunities for connecting with peers and communities beyond the immediate environment, e.g., peers from other parts of the country or world. Social media might also confer increased risk (Valkenburg et al., 2022), and some claim that the introduction of screen-use and social

media has contributed to the likely increase in mental health issues in the last decade, especially for girls (Twenge, 2020). One meta-analysis showed a dose-response relationship with a stronger association with depression with each hour spent on social media (Liu et al., 2022). However, most other meta-analyses show only modest associations with depression (Cunningham et al., 2021; Hancock et al., 2022; Vahedi & Zannella, 2021; Yoon et al., 2019), or mixed findings (Ivie et al., 2020). One recent TESS study (the same cohort as in the current work) found that the frequency of social media use did *not* predict depression across ages 10 to 16 (Steinsbekk et al., 2023), and Scherr et al. (2018) found that depression predicted social media use rather than the reverse. It is likely that whether social media is harmful depends on the activity being played out. For example, *cyberbullying* has been shown to predict depression (Hu et al., 2021). Moreover, Büttner and Rudert (2022) found that not being tagged in posts leads to an overall negative mood, especially in those scoring high on the need to belong—indicating potential detrimental effects of *social exclusion*. The research field is still relatively new, and future studies should keep investigating which activities may be detrimental and regarding which outcomes (e.g., not directly for depression, but sleep quality or body dissatisfaction). (See more in section 1.3.2.2. on body dissatisfaction).

Finally, the peer context may be an arena for co-rumination, which poses a risk for depression (Hankin et al., 2010; Spendelow et al., 2017; Stone et al., 2011). Co-rumination involves frequently and repeatedly discussing and speculating about the same problem, encouraging each other to discuss problems, and to focus on negative feelings (Rose, 2002). Co-rumination has potential advantages such as improved friendship quality (Felton et al., 2019)—perhaps because of the intimacy from social sharing (Rose, 2021). However, studies have also shown that co-rumination might generate interpersonal stress, which increases the risk of depression (Hankin et al., 2010; Rose et al., 2017). Examples are fighting with a friend or a romantic breakup (Hankin et al., 2010).

1.3.2.2. Relational self-esteem

Adolescence is marked by changes in self-esteem. Here, I will focus on *global* and *relational* self-esteem. Global self-esteem can be defined as a general evaluation of the self, and thus reflect how worthy the person feels (e.g., Rosenberg et al., 1995). The early adolescent experiences a drop in global self-esteem (Robins et al., 2002) which thereafter increases towards adulthood (von Soest et al., 2016).

Global self-esteem has also been defined as the sum of evaluations in domain-specific self-concepts (e.g., physical self-esteem, athletic competence) (Marsh & Shavelson, 1985). Harter et al. (1998) have proposed the term *relational self-esteem,* which entails that an individual's feeling of self-worth vary as a function of the relational context. This might be particularly relevant in the early adolescent period, marked by increased self-consciousness and concern with the evaluation of others (Harter, 2006; Westenberg et al., 2004). According to Harter (2015), self-esteem in different domains is related to how accepted and supported the early adolescent feels within different relationships: physical appearance, likability by peers; and athletic competence pertain to the peer context; academic competence and behavioral conduct pertain to the parent context. Studies have provided factor-analytic support for these domains in adolescents (e.g., Wichstrøm, 1995), and shown that physical appearance (body dissatisfaction) predicts lower global self-esteem (von Soest et al., 2016). Moreover, body dissatisfaction (Paxton et al., 2006; von Soest et al., 2016) and global self-esteem (e.g. Chang et al., 2018) both predict depression.

As reviewed above, social media represents a widening of the peer context and may conceivably influence the adolescent's relational self-esteem. However, findings are mixed, possibly depending on which activities adolescents engage in (reviewed by; Krause et al., 2021, and by; McLean et al., 2019). For example, based on a systematic review, Krause et al. (2021) conclude that social comparison may lead to lower self-esteem, and positive feedback

may lead to higher self-esteem. McLean et al. (2019) suggest that *social comparison* leads to body dissatisfaction specifically. This is in line with a TESS study showing that other-oriented—but not self-oriented—use of social media predicted more appearance dissatisfaction, at least in girls (Steinsbekk et al., 2021).

Evidently, the transition to adolescence is marked by important developmental advances involving puberty and psychosocial changes, which may render the early adolescent vulnerable to depression in various ways. Against this backdrop—I will now move on to the literature that more closely elucidates the research questions of this thesis: the stability of depression from preschool to adolescence (1.4.); the relationship between adolescent depression and personality (1.5.); and the female preponderance in adolescent depression (1.6.).

1.4. Does childhood depression increase the risk of adolescent depression?

Depression has been shown to be a recurrent disorder during adulthood (reviewed by; Richards, 2011), and from adolescence into adulthood (e.g., Johnson et al., 2018; Waszczuk et al., 2016). As previously described, depression may be identified in very young children (Vasileva et al., 2021). Even if it is low-prevalent, this opens the possibility that childhood depression increases the risk for adolescent depression. However, whether this is the case is uncertain. Moreover, as will be detailed below, different types of stability inform on this question.

1.4.1. The risk of staying a "high-scorer"

Previous research on the risk of recurrence has investigated the increased risk for future depression among those who have been depressed compared to those who have not: *rank-order stability*. Statistically, such stability has been investigated with odds ratios (OD), correlations, and predictions, which tap into an individual's risk of depression relative to the group. Most research has been done on adults (Richards, 2011) and from adolescence into

adulthood (e.g., Johnson et al., 2018). Even so, few studies have investigated younger samples.

Research on whether diagnosed (categorically defined) depression increases the risk of later depression may be interpreted as rank-order stability. Copeland et al. (2013a) did not identify rank-order stability in a combined measure of MDD and dysthymia diagnoses from late childhood into adolescence. However, given that depression was infrequent in late childhood in this sample, the study had low statistical power—rendering the results uncertain. Moreover, no study has investigated depression—as MDD and dysthymia are measured separately—in a community sample from preschool into adolescence.

Rank-order stability has also been investigated with a dimensional measurement approach. Two studies found evidence for rank-order stability (Finsaas et al., 2018; Mason et al., 2017). However, Mason et al. (2017) did not cover the age-span from childhood into adolescence, as they studied stability from early adolescence (age 11) and onwards. Furthermore, they used a self-reported questionnaire that only partly covered DSM-defined symptoms of MDD and dysthymia (an anxiety/depression sub-scale). Finsaas et al. (2018) studied rank-order stability from preschool into early adolescence (age 12) but did not cover the expected increase in depression *after* this age-period. Also, they used a combined measure of MDD, dysthymia, and depressive disorders not otherwise specified.

In sum, the magnitude of the rank-order stability of DSM-5-defined MDD and dysthymia from preschool to adolescence is undetermined. Also, MDD and dysthymia may show differing stability, but no studies have investigated these disorders separately nor their stability from preschool until *after* age 12-13 when depression first increases. Therefore, whether childhood MDD and dysthymia increase the risk of adolescent depression—when compared to other adolescents—needs further inquiry. This was one of the aims of Study I.

1.4.2. The risk of staying at the same level

Importantly, rank-order stability implies a risk for future depression when compared to *other* children: whether a child maintains their depression level relative to others. However, for clinical and preventative purposes, it is more valuable to know the likelihood that a depressed child or adolescent stays at the same level in the future—for example, to assess what may occur if the child does not receive treatment. *Absolute stability* concerns to what extent an individual at a specific depression level stays at this depression level in the future and can therefore not be illuminated by correlational approaches used for rank-order stability, which mix between-person and within-person information. As part of Study I, we investigated absolute stability across preschool into adolescence.

1.4.3. Does childhood depression leave "depressive scars?"

One possible explanation for the above-noted rank order stability is captured in the scar-hypothesis. The proposal is that depression changes characteristics of the child or adolescents and/or their environment on a relatively permanent basis (Rohde et al., 1994). These changed characteristics increase the risk of a new depressive episode (Rohde et al., 1994). A range of "scar"-mechanisms are possible. For example, depression may leave structural or functional changes at the brain level, such as reduced hippocampal volume (Chan et al., 2016; Mikolas et al., 2018), and reduced hippocampal volume may serve as a vulnerability for depression (Chen et al., 2010; Dedovic et al., 2010). Moreover, persistent social withdrawal may discourage others from interacting with the child or adolescent, and leave the child with less social support, isolation, and feelings of loneliness (Thompson et al., 2006)—all risk-factors for depression (Dunn & Sicouri, 2022; Thompson et al., 2006).

Alternatively—instead of depression leaving scars—the observed rank-order stability might be produced by stable effects from confounding third variables. As noted in Section 1.1., one type of confounding effect is time-invariant. These may stem, for example, from a

genetic vulnerability for depression, causing the depression to occur repeatedly (Wray et al., 2018). Within-person approaches that examine whether change in depression predicts future change using the person as their control—*stability of within-person change*—account for such stable confounding effects. One previous study, from the same data set as used in the present thesis, applied such methodology but with an age-range up to 10 years (Wichstrøm et al., 2017). Thus, it is still uncertain whether within-person increases or decreases predict later corresponding changes through childhood into adolescence. Stability of within-person changes was thus examined across ages 4 to 14 as part of Study I.

1.5. Adolescent depression and personality

Several integrative etiological models for depression include individual differences between adolescents' characteristic ways of thinking, feeling, and acting (e.g., Rudolph, 2009)—typically captured in the term *personality traits* (Caspi et al., 2005). Historically, research on adults has mainly focused on personality traits, while research on children has been concerned with *temperament*: individual differences in affect, activity, attention, and self-regulation (Rothbart, 2007). However, with increased recognition of the conceptual overlap between temperamental and personality traits in childhood (Caspi et al., 2005), it is now more common to investigate personality traits not only in adults but also in adolescents.

The most established model for personality traits is the Five Factor Model (FFM): Neuroticism, Extraversion, Conscientiousness, Agreeableness, and Openness (John et al., 2008). These traits—commonly named the Big Five—have been identified as early as in 10-year-olds (John et al., 2008; McCrae & Costa Jr, 2008; Soto et al., 2011). Neuroticism captures variations in susceptibility to general distress and negative emotions. In children, it is described as a tendency to be "fall apart" under stress, be guilt-prone, and be easily frightened. Low scores entail adaptability to novel situations, self-reliance, and self-confidence (Shiner & Caspi, 2003). Extraversion describes a tendency to be vigorously and

actively engaged with the world around them (Caspi & Shiner, 2008). Children scoring high on extraversion may be sociable, high-spirited, physically active, and outgoing, while introverted children may be more quiet and content to follow the lead of their peers (Caspi & Shiner, 2006; Shiner & Caspi, 2003). Conscientiousness captures differences in self-control (Caspi & Shiner, 2008). Children scoring high on conscientiousness are often described as attentive, planful, and thinking before acting, while those on the low end are careless, distractable, and quit easily (Caspi & Shiner, 2006). Agreeableness manifests as compassion, politeness, and trust in others (Soto & Tackett, 2015). Typical descriptors for children high on this trait are warm, empathic, generous, and kind—while children at the low end are described as stubborn, tease others, and are harder to manage for adults (Caspi & Shiner, 2006; Shiner & Caspi, 2003). Finally, openness to experience describes intellectual curiosity, creativity, and aesthetic sensitivity (Soto & Tackett, 2015). It may be expressed as a child's creativity and tendency to seek stimulation and actively explore new environments (Caspi et al., 2005). Of note, openness has received somewhat weaker empirical support in children than the other four traits (Caspi & Shiner, 2006).

1.5.1. Explanatory models

The Big Five personality traits have been found to correlate with depressive symptoms in adolescence (Klimstra et al., 2010). Different explanations for the relation between personality and psychopathology have been suggested (De Bolle et al., 2012; Nigg, 2006; Ormel et al., 2013; Tackett, 2006), and all may apply to adolescent depression and the Big Five personality traits. In the following sections, I will briefly describe these explanatory models. The main-focus will be on the predisposition and scar-explanations, as they represent the research questions in Study II: the reciprocal relation between depressive symptoms and the Big Five personality traits across ages 10 until 16.

The (1) *predisposition model* posits that particularly high or low levels of personality traits pose a vulnerability for psychopathology. Especially high or low levels of certain traits may increase the likelihood of other risk-factors—such as stressful life events—or moderate the effect of these same factors. For example, adolescents scoring low on extraversion tend to experience less social support (Swickert et al., 2002), which increases the risk for depression—especially when exposed to stressful life events (Rueger et al., 2016). Also, based on findings that many depressed adults show evidence of cognitive rigidity (Meiran et al., 2011) (i.e., low openness), Khoo and Simms (2018) suggest that people low on openness might be less adaptive and therefore more vulnerable to depression than those with a more flexible and broad view of the world (i.e., high openness). Finally, cross-sectional evidence on suggests that low conscientiousness may lead to problems in daily functioning (e.g., interpersonal), and, in turn to depressive symptoms (Naragon-Gainey & Simms, 2017).

According to (2) the *scar model*, depression may alter an individual's premorbid personality traits (Ormel et al., 2013). A potential scar-mechanism involves rumination. Research has shown an increased tendency to ruminate even after depression has subsided (Krause et al., 2018), and rumination has been associated with higher levels of neuroticism (Slavish et al., 2018). Another example may be related to social withdrawal, which is often part of a depressive episode. Social withdrawal can lead to lower friendship quality (Barzeva et al., 2022)—which ultimately may leave the adolescent less likely to engage socially with these friends after the depressive episode (i.e., lower scores on extraversion).

A variant of the predisposition model captures that personality traits change the *course* of psychopathology after its onset (a patho-plasty/exacerbation model; Klein et al., 2011; Ormel et al., 2013). This may influence the severity of depressive symptoms and response to treatment (Klein et al., 2011). A variant of the scar model is that the changes in personality traits due to depression disappear after the depressive episode has remitted (a state or

concomitant model; Klein et al., 2011; Ormel et al., 2013). Thus: traits are temporarily colored or distorted by the depressive episode and the individual's mood state (Klein et al., 2011).

The predisposition and scar models represent reciprocal causal relations between depression and personality. Alternatively, longitudinal associations between depression and personality may instead be caused by (3) *common causes*. This model proposes that psychopathology and personality traits have similar etiological influences (Klein et al., 2011; Ormel et al., 2013), for example, underlying genetic factors or family environment (Ejova et al., 2020). Finally, (4) the *spectrum model* states that personality traits and psychopathology are different manifestations of the same processes, with psychopathology representing extreme degrees of a personality trait (Ormel et al., 2013; Tackett, 2006). For example, high levels of neuroticism have been equated with depressive symptoms (Ormel et al., 2013).

1.5.2. Evidence for the predisposition and scar-explanations

The predisposition and scar explanations for the relation between adolescent depression and personality may be informed by research on the higher-order personality traits of the FFM. A meta-analysis on adult samples found that high neuroticism, low extraversion, and low conscientiousness predicted depression (Hakulinen et al., 2015)—in line with the predisposition model. Moreover, the study found that depression predicted changes in levels in *all* traits; higher neuroticism and lower extraversion, conscientiousness, agreeableness, and openness (Hakulinen et al., 2015). Thus, tentative support for the scar-model was even wider.

Of note, the prevalence of depression is relatively stable across adulthood (Richards, 2011), and adult personality is characterized by increased stability (Rantanen et al., 2007). By contrast, in the transition to adolescence, depression becomes more prevalent (Merikangas et al., 2010a). Also, across adolescence, there are considerable changes in personality traits (Soto & Tackett, 2015), such as a temporary decline in conscientiousness and agreeableness

in early adolescence, which thereafter inclines (Soto et al., 2011). Given the changes in depression and personality in this developmental period, the extent to which predisposition and scar models explain the association between depression and personality may differ in adults versus adolescents—and even *across* adolescence. In the following, I will review evidence that informs on the predisposition and scar-explanations in adolescents.

1.5.2.1. Personality traits conferring a vulnerability for depression in adolescence

Seven longitudinal studies on adolescents have investigated whether the Big Five personality traits predict depression, shedding light on the predisposition model. Six of these identified that higher levels of neuroticism predicted higher levels of depression (Calvete et al., 2016; Goldstein et al., 2020; Kercher et al., 2009; Klimstra et al., 2010; Yang et al., 2008; Zhang et al., 2020). The seventh study oversampled for adolescents with high levels of neuroticism and found that those who scored especially high on neuroticism evinced a *declining* trajectory of depression (Williams et al., 2021). Further, one study found that lower extraversion and conscientiousness scores predicted higher levels of depressive symptoms (Klimstra et al., 2010). However, others have found that neither extraversion (Calvete et al., 2016; Goldstein et al., 2018; Yang et al., 2008) nor conscientiousness (Goldstein et al., 2018) predicted depression. In contrast to Klimstra et al. (2010), two of these studies accounted for neuroticism (Goldstein et al., 2018; Yang et al., 2008). Finally, both Goldstein et al. (2018) and Klimstra et al. (2010) found that agreeableness and openness did not predict adolescent depression.

1.5.2.2. Does depression in adolescence impact personality development?

Only two studies inform on the scar-model. The first included all the Big Five traits and found that higher levels of depression predicted higher levels of neuroticism and lower degrees of extraversion, conscientiousness, and agreeableness (Klimstra et al., 2010). The

second study—which did not include analyses on the other four traits—found that higher levels of depression predicted lower levels of neuroticism (Zhang et al., 2020)

1.5.2.3. Knowledge gaps

Collectively: most studies investigating whether personality traits predict depression find that higher levels of neuroticism predict higher levels of depression. Evidence for extraversion and conscientiousness is mixed (Goldstein et al., 2018; Yang et al., 2008). Regarding whether depression predicts personality traits, the sole study investigating all traits found that higher levels of depression predicted higher levels of neuroticism and lower levels of extraversion, conscientiousness, and agreeableness (Klimstra et al., 2010). Some studies, but not all, have accounted for the effects of neuroticism. Moreover, prior studies have followed samples from ages 12 to 16 at baseline. One study investigated a sample from ages 10 to 11 (Zhang et al., 2020). However, the prevalence of depression first increases around age 12-13. To thoroughly inform on the development of adolescent depression, studies with younger baseline ages and follow-up measurements covering this age are needed. In Study II, we investigated predictions across the age-spans 10 to 12, 12 to 14 and 14 to 16—while accounting for the effect of neuroticism.

All previous research—including on adults (Hakulinen et al., 2015)—have applied traditional between-person methodology such as cross-lagged regression-based panel models (CLPMs). As described in Section 1.1., these analyses conflate between- and within-person information and leave the results open to the impact of time-invariant confounding effects. These confounding effects may stem from, for example, genetic factors exerting stable effects on depression and neuroticism over time (Adams et al., 2020; Kendler et al., 2019), or trait-like parenting factors—such as low parental monitoring—associated with both higher levels of depression (Hamza & Willoughby, 2011) and lower levels of conscientiousness (Smack et al., 2015). Study II was the first study to shed light on predisposition and scar-models by

using within-person methodology. That is: we investigated whether within-person changes in personality traits predicted within-person increases in depression—and vice versa.

With one exception (Goldstein et al., 2018; Goldstein et al., 2020), prior studies have relied on self-reported questionnaires of both depressive symptoms and personality traits. Clearly, using only one informant (the adolescent)—as well as the same method for measurement—increases the risk of common method bias (Klein et al., 2011). This may inflate the association between depression and personality. In Study II, we used multi-method assessments (semi-structured interviews of symptoms of depressive disorders and questionnaires for measuring personality) as well as multi-source assessments (adolescents and parents were interviewed on the adolescents' depressive symptoms)—thus minimizing common method bias.

Finally, as previously noted, the predisposition and scar explanations may differ across adolescence. In Study II, we were the first to investigate potential developmental differences in predictions between depression and personality across ages 10 to 16.

1.6. The female preponderance in adolescent depression

One of the most replicable findings in psychopathology is that more women than men suffer from depression (Kuehner, 2017). The female preponderance emerges by age 12-13 and reaches a ratio of 2:1 throughout adolescence, which endures for most of adulthood (Salk et al., 2017). The gender difference appears in population-based studies applying both clinical interviews and self-report-questionnaires (Salk et al., 2017; Shorey et al., 2022), across age-cohorts (Salk et al., 2017), Moreover, the gender difference in depression may even be widening, at least in adolescence (Bakken, 2020; Bråten et al., 2023; Platt et al., 2021).

Of note, some have suggested that the female preponderance in depression does not exist or is exaggerated. The *artifact hypothesis* proposes that because the society is more accepting of demonstrations of vulnerability in girls than boys, girls more readily express

their depression (Parker & Brotchie, 2010). Importantly, the artifact hypothesis is unresolved and deserves continued attention (Salk et al., 2017) (for a further discussion, see section 4.5.3.1.). The present work is based on the female preponderance as a factual finding and hence focuses on possible explanations. Despite decades of investigations, there is still a dearth of clear answers. In the following, I will briefly review the main explanatory models for the gender difference in depression, including the explanations that are the focus of Study III. In our study, gender was coded according to the child's biological sex at birth as registered in their national identification number. Of note, current theories and empirical studies concerning the gender difference in depression—including our study—treat gender as a dichotomous phenomenon, excluding questions concerning transgender and nonbinary individuals (Hyde et al., 2019) (for a further discussion on the gender binary see section 4.5.3.2.).

1.6.1. Explanatory models

Prior explanatory models are general elaborations of the classical stress-vulnerability model (Zubin & Spring, 1977). They encompass factors such as puberty, body dissatisfaction, and social challenges (Nolen-Hoeksema & Girgus, 1994), cognitive and genetic vulnerability, interpersonal stress (Hankin & Abramson, 2001), and pubertal hormones in interaction with gender role intensification (Cyranowski et al., 2000). To integrate these models, Hyde et al. (2008) developed the ABC-model (affective, biological, and cognitive factors), which proposes intricate interactions between different factors and that one factor may increase the likelihood of another, increasing an overall "depressogenic vulnerability". For example, Hyde and Mezulis (2020) suggest that because of biological and social changes, the adolescent becomes a self-critic of their appearance, which is more detrimental for girls than boys. In support of such a notion, research shows that self-surveillance (i.e., habitual monitoring of the body's appearance) predicts depressive symptoms more strongly for adolescent girls than

boys (Grabe et al., 2007). Overall, some elements of these prior models have gained empirical support, some partial support, and others have hardly been investigated (reviewed by; Hyde & Mezulis, 2020).

An important premise in prior models is a *stress exposure explanation* (Hammen, 2009a; Hankin et al., 2007): in early adolescence, girls become exposed to more stress, which partly explains why they also become more depressed. Several mechanisms may be possible to account for gendered stress exposure. As mentioned before, puberty brings a social reorientation towards peers, increased sensation-seeking and risk-taking behaviors (e.g., Bailen et al., 2019), as well as a peak in emotional reactivity and impulsivity (Rapee et al., 2019). Given that girls enter puberty one to two years prior to boys, these processes may ultimately contribute to girls becoming exposed to more psychosocial stress in the earliest adolescent years. Moreover, adolescence is marked by more prolonged negative affect—possibly to a stronger degree for girls than boys (Rapee et al., 2019)—setting the stage for more co-rumination (Rose et al., 2017). Research indicates that early adolescent girls engage in more co-rumination (Rose et al., 2017; Stone et al., 2011). Further on, co-rumination has been found to predict peer stress only in girls (Rose et al., 2017)—possibly contributing to a gender difference in stress exposure.

Another fundamental explanation for the female preponderance in depression is the *stress reactivity model* (Hammen, 2009a; Hankin et al., 2007): that girls react more strongly when exposed to stress in the form of greater levels of depression. Moreover, studies have shown that a heightened emotional reactivity in response to social cues tends to be higher in adolescent girls than boys (e.g., Bailen et al., 2019). This may help explain their increased vulnerability in this period characterized by vast social changes (i.e., including interpersonal stress). It is possible that early adolescent girls more often show dysregulated stress-responses in the body's main stress system, the hypothalamic-pituitary-adrenal (HPA) axis (Oldehinkel

& Bouma, 2011). This system is influenced by gonodal pubertal hormones (Roberts & Lopez-Duran, 2019)—implicating estradiol as a contributing factor in increased stress reactivity in girls.

1.6.2. Evidence for the stress exposure and reactivity explanations

The stress exposure and reactivity explanations are not mutually exclusive. Also, their validity may vary between stressors. In the work herein, the focus will be on two of the most important risk-factors for adolescent depression: stressful life events (SLEs) (Ge et al., 1994) and bullying victimization (Christina et al., 2021). Their potential involvement in the female preponderance in depression is undetermined and was therefore investigated in Study III—within a stress exposure and reactivity framework. We proposed specific criteria that should be fulfilled to indicate support for these explanations; criteria 1-3 are related to the exposure explanation, and criteria 4-5 to the reactivity explanation. In the following, I will review these criteria and relevant research.

1.6.2.1. Do adolescent girls become exposed to more SLEs and bullying victimization?

In Study III, we argued that three criteria must be met to indicate support for the stress exposure explanation for the female preponderance in depression. Criterion 1 states that girls become exposed to more stress than boys in the period just before the female preponderance in depression emerges (i.e., early adolescence), and not earlier (i.e., late childhood). Research on whether girls become exposed to more SLEs in adolescence than boys provides a mixed picture. Some studies show an increase in SLEs from childhood to adolescence (e.g., Larson & Ham, 1993), but it is undetermined whether this increase is stronger for girls. A meta-analysis indicates that girls are exposed to more SLEs than boys, and especially so in adolescence when compared to earlier ages (Davis et al., 1999). However, later studies have not found this gender difference (Jenness et al., 2019; Sund et al., 2003).

Whether girls become exposed to more bullying victimization in early adolescence (Criterion 1) is also uncertain. Most research has been conducted on gender differences in forms of bullying victimization. For example, one study found that in mid-adolescence, boys become exposed to more physical bullying and girls to more relational bullying (Hager & Leadbeater, 2016). Another study on early adolescents, however, did not find a gender difference in relational bullying (Lepore & Kliewer, 2019). One study found that girls were more exposed to cyberbullying than boys (Holfeld & Leadbeater, 2017), while another did not find this gender difference (Díaz & Fite, 2019)—both on early adolescent samples. Thus, evidence for specific forms of bullying in early adolescence is mixed. In Study III, we focused on bullying victimization overall. Research on "overall bullying load" indicate an increase only among girls (Wendelborg, 2020), or that girls and boys are exposed to an equal amount of bullying (Sweeting et al., 2006) during early adolescence.

Criterion 2 states that increased stress *before* early adolescence (that is: when the female preponderance in depression emerges) predicts increased depression in early adolescence—at least among the girls—and at the within-person level. Two previous studies investigated 13-year-olds and older: at the between-person level (Ge et al., 1994) and whether within-person changes in SLEs predicted the between-person level of depression (Jenness et al., 2019). Regarding bullying victimization, there is some within-person research, but only on the late childhood period (from age 9 until age 10; Kochel & Rafferty, 2020), and from age 12 and onwards (Davis et al., 2018)—that is: paralleling the age when the gender difference typically emerges (Salk et al., 2017). Summing up, there are no within-person studies on either SLEs or bullying victimization covering the age-period from late childhood into early adolescence.

Criterion 3 states that the emerging gender difference in depression should be explained by (i.e., accounted for or mediated) by a gender difference in stress exposure. To date, this has not been investigated for either SLEs or bulling victimization.

1.6.2.2. Do adolescent girls react stronger in the form of more depression?

To support a stress reactivity model, we argued that two criteria should be fulfilled. Criterion 4 states that increased stress should be more strongly associated with increased depression in girls than boys—and at the within-person level. According to Criterion 5, this gender difference in the strength of predictions from stress on depression should first appear in late childhood or early adolescence (i.e., just prior to the female preponderance in depression emerges)—not earlier childhood. Previous between-person research indicates that SLEs predict depression more strongly in girls than boys in early adolescence (Ge et al., 1994). Whether this is the case for bullying victimization is uncertain. Two studies found that bullying victimization predicted depressive symptoms among girls but not boys (Bond et al., 2001) or was stronger in girls than in boys (Lepore & Kliewer, 2019). A meta-analysis did not identify this gender difference in predictions from bullying victimization to internalizing disorders (i.e., anxiety and depression measured collectively) (Christina et al., 2021). However, the study did not investigate if the strength of predictions to *depression* varied by gender. Summing up, we lack research on whether predictions from both SLEs and bullying victimization to depression are stronger among girls than boys—and at the within-person level (Criterion 4). We also lack research on whether stronger predictions for girls are specific to the period just before the female preponderance in depression emerges and not at younger ages (Criterion 5).

1.6.2.3. Knowledge gaps

Collectively, evidence for the five criteria needed to support the stress exposure and reactivity models is uncertain. First and foremost, there is a lack of studies that cover—and

contrast—the age-period from before and after the gender difference emerges, which is necessary for several of the criteria. Whether SLEs and bullying victimization increase more for girls in the transition to adolescence (Criterion 1) is undetermined (Davis et al., 1999; Jenness et al., 2019; Sund et al., 2003). No studies have investigated whether these stressors predict depression for early adolescent girls at the within-person level (Criterion 2), and finally: whether increased exposure mediates the gender difference in depression (Criterion 3). Furthermore, whether increased stress is more strongly related to depression in girls than boys at the within-person level (Criterion 4) and contrasted to the age-period before the gender difference emerges (Criterion 5) has not been investigated. To shed light on the stress exposure and reactivity explanations, these five criteria were investigated in Study III across ages 8 to 14.

Importantly, most prior research on SLEs and bullying victimization has utilized traditional between-person methods. Again, as described previously, such methods do not separate between- and within-person information. Thus, they are influenced by time-invariant confounding effects, stemming, for example, from genes involved in both SLEs and depression (Clarke et al., 2018) or persistently family conflict that provide an increased risk for both bullying victimization (Hemphill et al., 2015) and depression (Buehler, 2020). In Study III, we therefore specified (and examined) predictions between stressors and depression at the within-person level.

Finally, all prior studies have used self-reported questionnaires to measure depression, SLEs, and bullying victimization, which increases the risk for common-methods effects. In Study III, we minimized such bias by using multi-method assessments (interviews for symptoms of depressive disorders and questionnaires for SLEs and bullying victimization) as well as multi-source assessments (parents and adolescents were interviewed about depressive symptoms and reported on SLEs, and teachers reported on bullying victimization).

1.7. Aims of the Thesis

This thesis aspired to extend previous research on the development of child and adolescent depression by using the screen-stratified sample of the Trondheim Early Secure Study (TESS) and within-person longitudinal methodology. Symptoms of DSM-defined MDD and dysthymia were measured biannually measured with semi-structured interviews. Data from waves T1-T7 was applied (Study I: ages 4-14, Study II: ages 10-16; Study III: ages 8-14).

- The overall aim of Study I was to investigate different types of stability in MDD and dysthymia across preschool, childhood, and into adolescence. We investigated a) stability of form with i) partial measurement invariance; to what extent factor loadings from a latent depression construct loads similarly on each of the nine symptoms specified for MDD and seven symptoms specified for dysthymia; and ii) the frequency of different symptoms across ages. We also investigated d) the prevalence of MDD and dysthymia (both symptoms and disorders). Finally, we investigated a) rank-order stability, b) absolute stability, and c) within-person stability of change (i.e., whether within-person decreases or increases in depression predicted corresponding changes in depression two and four years later).
- The aim of Study II was to illuminate the potential reciprocal relationship between adolescent depression and personality traits. We investigated whether within-person changes in personality traits predicted within-person increases in depressive symptoms (a collective measure of symptoms of MDD and dysthymia), to inform on the predisposition explanation. We also investigated whether within-person increases in depressive symptoms predicted within-person changes in personality traits—in line with a scar explanation. Lastly, we tested for potential developmental differences—whether predictions differed across ages 10 to 16.

- The aim of Study III was to shed light on possible explanations for the female preponderance in depression emerging in early adolescence. We investigated whether girls became exposed to more SLEs and bullying victimization prior to the emerging gender difference in depression. Moreover, whether gendered stress exposure mediated the female preponderance in depressive symptoms (a collective measure of symptoms of MDD and dysthymia). These findings would be in line with a stress exposure explanation. Secondly, we investigated whether within-person increases in these stressors predicted within-person increases in depressive symptoms more so for girls than boys—in line with a stress reactivity explanation.

2. METHODS

2.1.Study Design and Procedure

The Trondheim Early Secure Study (TESS) is an ongoing prospective study from preschool to early adulthood (Steinsbekk & Wichstrøm, 2018). The overall aim is to investigate mental health, health behavior, and psychosocial development in children, adolescents, and young adults. Information on a wide range of risk and vulnerability factors of importance for child development is collected from the child, parents, and teachers through observations, tests, clinical interviews, and questionnaires. Two birth cohorts (N = 3,456) in Trondheim—the third most populated city in Norway at the time—were enrolled at age 4 (in 2007 and 2008) and have been examined biennially thereafter. At the time of recruitment, the population of Trondheim was representative of the national population of Norway on several key parameters (e.g., employment rate, educational levels, gross income; Statistics Norway, 2012). Study I on stability in depression was based on data from the first six data waves (ages 4, 6, 8, 10, 12, and 14). Study II on depression and personality in adolescence was based on four data waves (ages 10, 12, 14, and 16). Study III on the female preponderance in depression was also based on data from four data waves (ages 8, 10, 12, and 14).

2.2.Recruitment and Screening

In Norway, health check-ups for all children are organized by community health-child clinics[3] at regular intervals from a child is born until age 4. Ahead of the health check-up for 4-year-olds, parents of the children in two birth-cohorts (born in 2003 and 2004) received a letter with information regarding the TESS and an invitation to participate. Additionally, they received a screening assessment questionnaire for emotional and behavioral problems—the Strengths and Difficulties Questionnaire (SDQ; 31 items) version 4–16 (Goodman et al., 2000). The SDQ has been shown to efficiently screen for emotional and behavioral problems in 4-year-olds (Sveen et al., 2013). Parents were encouraged to complete the SDQ and bring it to the health-check-up. SDQ-values were used in the stratification of the sample (described in section 2.3.).

At the 14 well-child clinics in Trondheim, almost all children, accompanied by their parents (n = 3,358, 97.2%), attended. At the health checkup the parents received further information on the TESS from a health nurse, in accordance with procedures approved by the Regional Committee for Medical and Health Research Ethics (REK), Mid-Norway. Parents with insufficient proficiency in Norwegian to complete the SDQ were excluded (n = 176). 166 parents were mistakenly not asked to participate by the health nurses. Among those asked to participate (n = 3,016), 2,477 (82.1%) accepted and gave written consent, while 539 declined (17.9%). More details on the procedure and participation rates are displayed in Appendix Figure I: Flowchart of Sample Recruitment.

2.3.Stratification

With 97.2% of the population attending the well-child clinics in Norway, the TESS sample is regarded as a community sample (Steinsbekk & Wichstrøm, 2018). The overall aim of the TESS was to investigate factors that may influence mental health and psychosocial

[3] https://www.helsenorge.no/hjelpetilbud-i-kommunene/helsestasjon-0-5-ar/firearskontroll/

development. Therefore, based on information from the SDQ, children with emotional and behavioral problems were oversampled to increase the sample variability and thus statistical power. The probability of being drawn to participate increased with a higher score on total difficulties on the SDQ (20 items, scored 0-2, assessing emotional symptoms, and conduct and peer relationship problems). By using a random number generator, a total of 1,250 children were drawn to participate. Of these, 1,007 (80.6%) participated at baseline (T1). As shown in Appendix Figure II, the children were divided into the following four strata based on their SDQ total difficulties score: 0-4 (44.2% of the population), 5-8 (29.5%), 9-11 (18.5%), and 12-40 (7.8%). The drawing probabilities within the respective strata were 37%, 48%, 70% and 89%. This oversampling of mental health problems was accounted for in the analyses. The drop-out rate after consent at the well-child clinic did not differ across the four SDQ strata; (3)=5.70, p=.127, or by gender; (3)=0.23, p=.973.

2.4.Participants

1,250 children were randomly selected for participation in the study. Of these, 1,007 were successfully enrolled at Time 1. The sample included 49.9% boys, and the mean age at each wave was as follows: T1: M_{age}=4.59, SD=0.25; T2: M_{age}=6.72, SD=0.19; T3: M_{age}=8.79, SD=0.23; T4: M_{age}=10.51, SD=0.17; T5: M_{age}=12.50, SD=0.14; T6: M_{age}=14.35, SD=0.14; T7: M_{age}=16.98, SD=0.31. Participants with information from at least one wave composed the analytical sample (Study I: n = 1,042, Study II: n = 817, Study III: n = 748).

The sample, adjusted for stratification, is representative of the Norwegian population regarding parents' level of education (Statistics Norway, 2012) and family variables (The Norwegian Directorate for Children, 2017). The one exception was a higher divorce rate in our sample (7.6%) when compared to the general population (2.1%) (Statistics Norway, 2017). The differences in rates of occupational categories between the sample and Trondheim were negligible, below 3.6% (Steinsbekk & Wichstrøm, 2018). Unweighted sample

characteristics (n = 1,007) at baseline are displayed in Table 2 (the table is reported as Table 1 in Study II).

Table 2 Sample Characteristics at T1

Characteristics		%
Gender of child	Male	48.9
	Female	51.1
Gender of parent informant	Male	16.7
	Female	83.3
Parent informant	Biological parent	98.3
	Adoptive parent	1.3
	Foster parent	0.4
Biological parents' marital status	Married	59.3
	Cohabitating > 6 months	21.9
	Cohabitating < 6 months	0.4
	Divorced/separated/no longer cohabitating	16.4
	Widowed	0.1
	Never lived together	1.9
Ethnic origin of biological mother	Norwegian	93.0
	Western Countries	2.7
	Other Countries	4.3
Ethnic origin of biological father	Norwegian	91.0
	Western Countries	5.8
	Other Countries	3.2
Informant parents' socioeconomic status	Leader	17.5
	Professional, higher level	30.1
	Professional, lower level	30.1
	Formally skilled worker	18.5
	Farmer/fishermen	0.2
	Unskilled worker	3.6
Parent's highest completed education	Did not complete junior high school	0.0
	Junior high school (10th grade	0.6
	Some education after junior high school	6.1
	Some collage- or university education	7.6
	Bachelor's degree	6.2
	College degree (3-4 years study)	20.3
	Master's degree or similar	20.3
	PhD completed or ongoing	4.4

2.5.Data Collection Procedure

A few weeks after the well-child clinic visit, the parent and child were invited to the university clinic at the Norwegian University of Science and Technology for testing and interviews. The families were assigned two research assistants (one dedicated to the child, the other to the parent). The research assistants (n = 7) had (i) at least a bachelor's degree in a relevant field, (ii) extensive experience working with children and families, and (iii) substantial training in diagnostic interviewing and coding. The test day normally lasted 4-5 hours, and involved the completion of questionnaires, interviews, and parent-child cooperative tasks. Completed questionnaires from the daycare or schoolteacher who knew the child best were obtained in the weeks after the test day if the parents had given permission to collect such information.

The test day procedure described here was repeated biennially from September 2007 to January 2022 (T1-T7), although the content of the testing differed (e.g., developmentally appropriate measures). When possible, the same research assistants met with the family at every assessment point. Notably, T7 (from 28th of January 2020 until 1st of January 2022) was collected during the COVID-19 pandemic. However, as in previous waves, the data-collection was performed through face-to-face-interviews and self-reported questionnaires. The only exception was that we omitted the electrophysiology measurements (to measure autonomic reactivity), which had no relevance for the current work. Before the test-day, the participants were sent an information sheet on particulars regarding COVID-19. This included a reminder to be aware of symptoms before attending; the possibility of postponing or not attending for this or other reasons, and a reminder of common infection control measures. On the test-day, additional procedures were set in place to limit the risk of infection. These procedures were based on a risk-analysis performed by the Occupational Health Service at the Norwegian University of Science and Technology. For example, when the participants arrived, the

research assistants provided information on the infection control measures currently recommended by the Norwegian Institute of Public Health; used masks when this was recommended; were responsible for keeping at least two meters distance; and asked the participant whether they lived with someone with COVID-19 to assure they kept the current national quarantine rules. Summing up, despite the pandemic, we were able to collect data on almost all study variables, including depressive symptoms. However, as can be seen from the age at the 7th data collection wave (participants were on average almost 17 years of age), the collection was at times delayed due to quarantines at schools, potential or actual infection among the participating families, or general lockdowns in Trondheim or Norway.

2.6.Attrition rate

Those who consented to participate at the well-child clinic (and drawn to participate)—but thereafter did not attend at baseline—did not differ from those who attended regarding SDQ-strata group classification or gender (Steinsbekk & Wichstrøm, 2018). After an initial drop in retention from T1 to T2, there is generally low attrition in the TESS (Steinsbekk & Wichstrøm, 2018). Except for those who withdrew from the study, families were reinvited at later measurement points. Families not responding to the invitation on two consecutive occasions were not asked again, nor were those who withdrew from the study. Further, for the analysis of systematic attrition, attrition was defined as not participating in the data wave in question. Predictors of attrition were therefore examined by regressing attrition on every study variable at the preceding time points. In Study I, depressive symptoms at some waves predicted reduced retention at later waves: Symptoms of MDD at age 6 predicted attrition at age 8 (OR=0.82, 95% CI [0.67, 1.00]) and symptoms of MDD at age 12 predicted attrition at age 14 (OR=0.72, CI [0.59, 0.87]). Symptoms of dysthymia at age 4 predicted attrition at age 12 (OR=0.84, CI [0.70, 1.00]); symptoms of dysthymia at age 6 predicted attrition at ages 10 (OR=0.78, CI [0.63, 0.97]) and 12 (OR=0.80, CI [0.65, 0.98]); and finally,

symptoms of dysthymia at age 12 predicted attrition at age 14 (OR=0.74, CI [0.61, 0.89]). In Studies II and III, at age 12 symptoms of MDD (OR=1.39, CI [1.15, 1.70]) and dysthymia (OR=1.35, CI [1.12, 1.64]) predicted attrition at age 14. In Study III, bullying victimization at age 6 predicted attrition at ages 10, 12 and 14 (all ORs=1.02, CI [1.01, 1.03]).

Even though the abovementioned analyses suggested selective attrition according to study variables, many attrition tests were run, which increased the risk of false discovery of predicting variables. Therefore, an overall Little Missing Completely at Random (MCAR) test (Little, 1988) was conducted in Studies II and III. The results confirmed that the data were neither MCAR in Study II (χ^2=1286.46, df=935, p<.001) nor in Study III (χ^2=256.01, df=220, p=.048). However, the normed tests were 1.38 (Study II) and 1.16 (Study III), which is below the suggested cut-off of 2 (Ullman, 2001), indicating that data were at least missing at random (MAR).

2.7.Measures

2.7.1. Depression

Symptoms of MDD and dysthymia were measured by The Preschool Age Psychiatric Assessment (PAPA; Egger & Angold, 2006), the Child and Adolescent Psychiatric Assessment (CAPA; Angold & Costello, 2000), and the Schedule for Affective Disorders and Schizophrenia for School-Age Children (Kiddie-SADS; Kaufman et al., 2016). The PAPA is a parent-reported, developmentally appropriate, semi-structured diagnostic interview for children aged 2 to 7 years and was therefore administered when the children were 4 (T1/baseline) and 6 years old (T2). The PAPA is a revised version of the CAPA. The CAPA is appropriate for interviewing older children and their parents, and was used at ages 8, 10, 12, and 14 (T3-T6). At age 16 (T7), symptoms of MDD and dysthymia were measured by the Kiddie-SADS. Note that from age 8 and beyond, the children and the parents were interviewed separately—hence providing two sources of information on depressive

symptoms. A symptom was considered present if reported by either a child or a parent, and symptom count scores were created as the sum of MDD and dysthymia symptoms. MDD and dysthymia were analyzed separately in Study I and collectively in Studies II and III. As described in Section 1.2.1.1., in the PAPA, persistent preoccupation with play themes involving death or self-harm was added as an age-appropriate criterion for suicidality (Luby et al., 2002).

Interviewers had been trained with the team that developed the PAPA/CAPA, as well as with clinical child psychologists or a child psychiatrist, when learning how to administer and score the Kiddie-SADS. Interrater reliabilities were estimated with intraclass correlations (ICC). At age 4 PAPA interviews, interrater reliabilities among blinded coders of 9% of videotapes were ICC=.91 for symptoms of MDD and ICC=.89 for symptoms of dysthymia. At age 10 CAPA interviews, inter-rater reliabilities among blinded coders of 15% of audiotapes were ICC=.87 for symptoms of MDD and ICC=.85 for symptoms of dysthymia. At age 16 Kiddie-SADS interviews, interrater reliabilities among blinded coders of 17% of audiotapes were ICC=.81 for symptoms of MDD and ICC=.76 for symptoms of dysthymia.

2.7.2. Personality Traits

In Study II, personality traits were measured across ages 10 to 16 by the Norwegian version of the self-reported Big Five Inventory (BFI-46A; Soto et al., 2008). The BFI consists of 46 items (44 used in the analysis, two are optional) assessing Neuroticism (8 items), Extraversion (8 items), Conscientiousness (9 items), Agreeableness (9 items) and Openness (10 items). Response options range from 1 (*disagree*) to 5 (*agree*). The mean scores of the five traits were applied. The internal consistency was as follows, from ages 10 to 16 respectively: Neuroticism: α=.59, .72, .81, .83; Extraversion: α=.54, .67, .75, .81; Conscientiousness: α=.65, .72, .77, .76; Agreeableness: α=.64, .71, .72, .71; Openness: α=.69, .74, .74, .76.

2.7.3. Stressful Life Events

SLEs were assessed in Study III at ages 8, 10, 12, and 14 with a parent and child self-reported checklist covering 31 SLEs occurring since the previous measurement wave. The SLEs varied from important but not life-threatening life events (e.g., change of school, a step-parent moving into the home) to less common but very serious ones (e.g., sexual abuse) (see Appendix for the complete list). A SLE was considered present if reported by either child or parent. A total score was created as the sum of the number of SLEs. When the children completed the questionnaire (from age 8), they were encouraged by the research assistant to ask questions about items they did not understand.

Given the very different degree of seriousness of the events the children and parents reported on, we tested whether predictions between SLEs and depressive symptoms were driven by the more serious life events. After all, effects could likely be driven by the most adverse events. We therefore compared the correlations between depressive symptoms and important (but not life-threatening) life events to the those between depressive symptoms and events with a substantial potential for grave physical and mental harm. Here we used the Satorra-Bentler scaled chi-square difference test (Satorra & Bentler, 2001). Allowing the correlations to be different did, however, not improve model fit when compared to when the correlations were set to be identical ($\Delta\chi^2(4)=2.49$, $p=.952$). This suggests that the relationship between SLEs and depression did not differ according to the seriousness of the SLEs.

2.7.4. Bullying Victimization

Bullying victimization was measured in Study III across ages 8–14 by a teacher version of the Olweus Bully Victim Questionnaire (OBVQ; Solberg & Olweus, 2003). The questionnaire was completed by the TESS participant's primary teacher. The OBVQ teacher version consists of five items pertaining to both physical bullying and social exclusion during the last 3 months. The items tap into the frequency of physical harm, verbal abuse, social

exclusion, being overlooked, and belongings being hidden or destroyed. Response options range from *Never*, *Rarely*, *1-3 times per month*, *1-4 times per week* to *every day*. We constructed a sum score representing the overall number of bullying victimizations during the previous 3-month period using mid-point values (e.g., 1-3 timers per month was coded as 6 times the last 3 months). The internal consistency was as follows at ages 8, 10, 12 and 14 respectively: α=.77, .77, .79, .69.

2.7.5. Sociodemographic information

Sociodemographic information on the child and parent was reported by the parent during the diagnostic interviews (PAPA and CAPA). Throughout the thesis and in line with what we did in the three papers, I adhere to the Publication Manual of the American Psychological Association, 7th edition (APA-7; American Psychological Association, 2020). According to APA-7, the term "female" should only be used as an adjective. Thus, I avoid the terms "males" and "females" as nouns and instead use "boy" and "girl". As noted in Section 1.3.2., while we used the terms "him" or "her" in the papers, in the thesis I use the term "they/themselves/theirs".

2.8.Ethics

In Norway, research covered by the Health Research Act[4] of 2009 as well as the Research Ethics Act[5] of 2017 must be approved by the REK in any given region. REK is founded in Norwegian law and research ethics, and international conventions such as the Declaration of Helskinki[6]. The Norwegian Data Protection Authority (NDPA) was responsible for the judgement of data-protection before 2009, REK was responsible

[4] https://lovdata.no/dokument/NL/lov/2008-06-20-44
[5] https://lovdata.no/dokument/NL/lov/2017-04-28-23
[6] https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medicalresearch-involving-human-subjects/

thereafter. Thus, the initiation of TESS in 2007 (T1) was approved by both NDPA and REK Mid-Norway (approval number 2009/994).

All parents of participating children provided consent on their behalf before reached 16 years of age. At age 16, the adolescents consented themselves. In Norway, when a child is below age 16, their parents decide whether they may participate in research. However, according to the Children Act[7] (§33) the child's opinion should be taken into consideration—and increasingly so as the child grows older—both by the parents and the research workers. In the TESS, at age 12 the children were specifically informed about the study. Moreover, across all ages, the research assistants would pay attention to the child's or adolescent's motivation and comfort. They would as much as possible ascertain that the child or adolescent did not participate in the various tests and observations if they did not want to. Also, research ethics was discussed regularly in the research group's meetings (every second week).

The TESS research assistants focused on the well-being and comfort of the children or adolescents through the various tasks they performed on the test day. If the research assistants became concerned about the participants' wellbeing, they were obliged to discuss this with the principal investigator—who is specialized in clinical psychology—to assess the need for taking further steps. If the participant/parent experienced emotional difficulties due to the data collection, they were offered a consultation with a clinical psychologist not involved in the TESS. This was not requested by any of the participants. The clinical interviews and most of the test day took place in soundproof rooms with one-way mirrors and built-in video recording systems at the university clinic. All video tapes and data files were saved on their own server on a local network, only accessible to the research assistants conducting the data collection, the project manager, and IT-personnel.

[7] https://lovdata.no/dokument/NL/lov/1981-04-08-7#KAPITTEL_6

In accordance with REK standards, no notable regular reward was given for the TESS participation, but children/adolescents received a small gift at the end of the day (e.g., water bottle, power bank). Parents received a 300 NOK gift-card (about 28 US$) and compensation for travel expenses or parking fees. There was, however, a family travel gift-card lottery prize of 40,000 NOK (about 3,800 US$) among participating families after each measurement point.

2.9.Statistical Analyses

All analyses were performed in Mplus (Study I: version 8.1, Study II and III: version 8.5) using a robust maximum likelihood estimator and probability weights to correct for the oversampling of children with mental health problems. Missing data were handled using a full information maximum likelihood (FIML) procedure under the assumption that the data was MAR (see also section 2.6.). We used within-person models to examine longitudinal associations: a random-intercept cross-lagged panel model (RI-CLPM; Hamaker et al., 2015) and autoregressive latent trajectory models with structured residuals (ALT-SR; Berry & Willoughby, 2017). As previously described, these methods model within-person development while accounting for time-invariant confounding effects from unmeasured variables.

The within-person statistical model applied varied between studies. In Studies I and II we used RI-CLPM. In this model, each person's deviation from her or his mean during the observation period is captured, and the factor loadings are set to be equal across time (that is, 1). Hence, no change in the importance of the variables to the latent random intercept across time is modeled. Thus, if there is a normative or common change in the contribution of a phenomenon to the overall level (random intercept) across development, this will not be captured. The advantage is a complete separation of between- and within-person information, and the time-varying latent variables capture deviation from the person's overall mean. Of

note, in our case, it was expected that participants would evince an increase in depression from their overall mean in early and middle adolescence, given that depression was expected to become more prevalent at these ages. In Study III, we used ALT-SR. This model captures both a person's deviation from their mean (intercept), and from their expected linear change through development (Berry & Willoughby, 2017). However, it is also possible to model deviations from a non-linear development—which we did in Study III because the change in depression level over time was not expected to be linear.

2.9.1. Study I

In Study I, we examined five different forms of stability in depression across preschool into adolescence (ages 4 to 14). (a) Stability of form was first examined as i) structural stability (partial measurement invariance) by Confirmatory Factor Analysis (CFA), comparing the model fit when all factor loadings were freely estimated to when they were fixed to be equivalent to two adjoining ages (Widaman et al., 2010). Due to our variables being categorical, chi-squares were not provided, and the Akaike information criterion (AIC) ≥2 was therefore used as the criterion (Burnham & Anderson, 2004). ii) The frequency of symptoms was examined with percentages by age group, and linear and quadratic latent growth curves.

(b) Prevalence, or group-level stability, was examined in two steps. First, we inspected percentages receiving the diagnosis of MDD and dysthymia by mean levels of symptoms. Thereafter, changes in the number of symptoms were examined by latent growth curves, considering symptom counts to follow a negative binominal distribution.

(c) Rank-order stability, or stability relative to the group (i.e., whether the children maintain their level of depression relative to the group) was investigated with Pearson's r. Furthermore, to test whether stability increased or decreased with age, we compared models

where correlations between adjacent ages were fixed to be equivalent as opposed to freely estimated, using the Satorra–Bentler scaled chi-square test.

(d) Absolute stability (i.e., whether children maintain their depression level over time)—or stability relative to oneself—was investigated with ICC. ICC captures the proportion of the total variance attributed to between-person variance; $= \frac{\text{between}}{\text{between+within}}$. In other words: if individual children typically stayed at their "absolute" depression levels across ages 4 to 14 (T1-T6), within-person variation would be low (closer to 0) and ICC closer to 1.

Finally, (e) stability of within-person changes was examined in a model drawing on the RI-CLPM, separating between-person differences from within-person changes (Hamaker et al., 2015), but here only examining autoregressions. As noted in Section 2.8.1., intercepts (in depression levels) were allowed to vary. Variations from the initial mean level in depression are therefore estimates of deviations from the child's own mean depression level across time. Thus, the model estimated whether deviation of one's mean level of depression predicted later deviations, while ruling out time-invariant confounding effects. In the following, I will describe this procedure as applied in Study I. In the RI-CLPM, between-person time-invariant effects were captured by creating a latent variable (a random intercept factor) for the depression construct, by setting factor loadings across ages 4 to 14 to 1. This allows the between-person stability—that is: differences between individual depression scores co-varying across all these time-points—to be transferred to this overall random intercept. Within-person lagged effects were modeled by creating a latent variable for each observed depression score—in this case, scores across ages 4, 6, 8, 10, 12 and 14. These latent variables were set with a factor loading at 1 and the variance of the observed score at 0. This transfers the variance of the observed scores to the latent variables and represents the deviations from that child's average depression level across age 4 to 14. Latent variables thus represent changes in depression from the individual mean depression level across time. These latent

variables at a certain timepoint were regressed on the latent variables two years earlier, from T2 (age 6) to T1 (age 4)—and two and four years earlier in the older age-groups (e.g., from age 8 on both ages 6 and 4). The RI-CLPM for Study I is depicted in Figure 1.

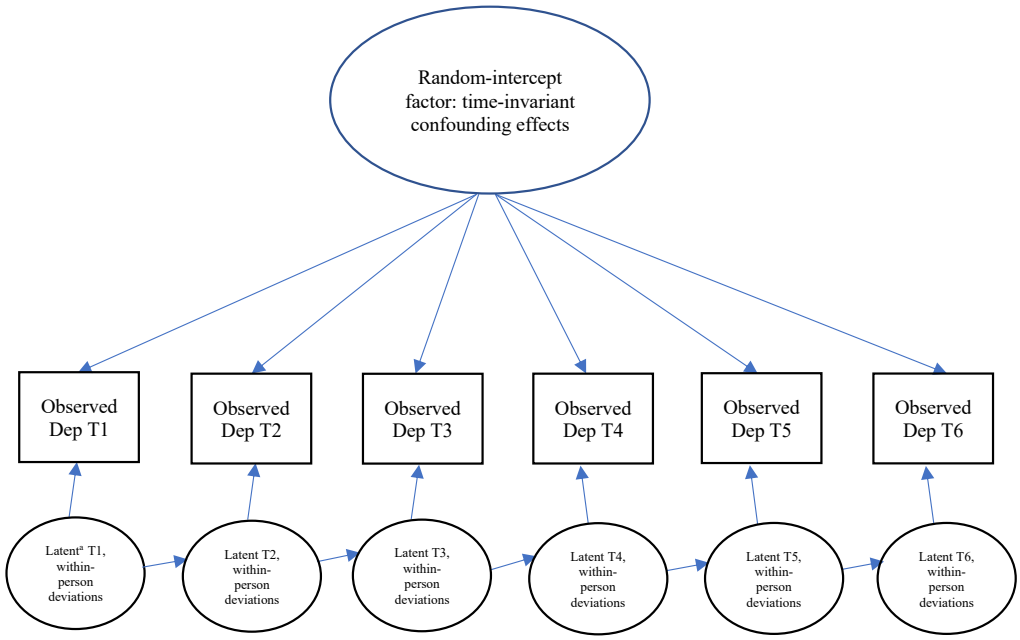


Figure 1: RI-CLPM for depression across ages

Note. In Study I we investigated predictions across two-year age-span—as depicted in the figure—but also four-year-age spans (e.g., T1 to T3). a = Latent variable.

Study II

In Study II, to illuminate the predisposition and scar explanations for the associations between adolescent depression and personality, we examined reciprocal predictions between depressive symptoms and the Big Five personality traits across ages 10 to 16. Prospective relations were investigated with RI-CLPM models. Within-person development was modeled for depression as in Study I (section 2.9.1.), and for each of the five personality traits. Thus, at each wave, the observed depression and personality trait scores were decomposed into

between-person stable and within-person varying parts. One random intercept factor for depression and one for the personality trait in question captured the overall levels of the two constructs. Latent depression and personality trait variables at each time point captured the adolescent's deviation from her or his own mean score across time. These latent deviations at one time-point were regressed on the latent changes at the time-point two years earlier (age 12 at age 10, age 14 at age 12, and age 16 at age 14). Deviations from mean levels of personality traits predicting later deviations in depression would be in line with the predisposition explanation. Deviations from mean levels of depressive symptoms predicting later deviations in personality traits would be in line with the scar explanation. With RI-CLPMs, time-invariant confounding effects underlying both depressive symptoms and personality traits were accounted for.

Of note, the RI-CLPM is power-demanding. Our sample size (817) was somewhat lower than the recommended size (e.g., Masselink et al., 2018). Therefore, we examined depressive symptoms and the Big Five personality traits in five separate models. However, previous studies examining several traits in multivariate models have found only neuroticism to predict depression (Goldstein et al., 2018; Yang et al., 2008). Moreover, neuroticism is correlated with the other personality traits (Van der Linden et al., 2010). This indicates the importance of controlling for this trait. Therefore, when examining extraversion, conscientiousness, agreeableness, and openness, we adjusted for neuroticism. Finally, in early adolescence, in addition to the female preponderance in depression (Salk et al., 2017), a study on adolescents found an increase in neuroticism for girls and a decrease for boys (Soto et al., 2011). Gender may thus have time-varying effects on both depression and personality. Again, given our sample size, we were not positioned to analyze girls and boys in separate models. We did, however, include gender as an observed covariate due to its potential time-varying confounding effects (Mulder & Hamaker, 2020).

Our secondary aim in Study II was to check for potential developmental differences. We did this by testing if a model where the cross-lagged paths were set to be equal across all ages fitted the data worse than a model where the paths between depression and personality traits were freely estimated (i.e., allowing the effects to be unequal across ages). If a model where cross-lagged paths were set to be equal did not deteriorate the model fit, we would prefer such a constrained model for parsimonious reasons. Such a result would indicate no developmental differences across these paths.

To compare our results with models in which neuroticism and gender were not accounted for, the RI-CLPM models were also run without including these factors as covariates. Finally, we also reran our models with traditional CLPM methodology to ease comparison with results from previous between-person studies (e.g., not accounting for time-invariant confounding effects).

2.9.2. Study III

In Study III, the overall aim was to inform on the stress exposure and stress reactivity explanations for the emerging female preponderance in depression. The stressors we examined were SLEs and bullying victimization. We suggested and investigated five criteria: three pertaining to the stress exposure explanation and two pertaining to the stress reactivity explanation. These criteria entail the need to include baseline ages younger than when the gender difference emerges around ages 12 to 13. We therefore included the measurement waves from ages 8 to 14 (T3-T6).

Criterion 1—whether girls become more exposed to stress than boys just prior to the emergence of the female preponderance in depression—and not earlier on—was examined in two steps. First, we inspected whether latent change scores in SLEs and bullying victimization increased among girls, specifically from ages 10 to 12 and not 8 to 10. Second,

we examined whether the potential increases in SLEs or bullying victimization from age 10 to age 12 were predicted by gender.

We tested Criterion 2—whether increased SLEs and bullying victimization predicted later depressive symptoms in girls at the within-person level—with a modified version of the ALT-SR model (decipted in Study III, Figure 1). In this model, between-person differences in a construct are captured by the intercept (representing the overall mean level), allowing the initial within-person levels of a construct (the intercepts) to vary—which is also the case in the RI-CLPM. ALT-SR additionally allows the slopes to vary. As such, between-person differences in depressive symptoms, SLEs, and bullying victimization were captured by the intercept (as in the RI-CLPM), but also the slope (representing the overall growth) pertaining to each of these three constructs. Within-person scores at each timepoint provide information about a person's deviation from their mean intercepts and slopes across ages 8 to 14. Of note, we expected an increase in depression across the included ages and potentially also in SLEs and bullying victimization (i.e., nonlinear developments). We therefore applied an ALT-SR model where the growth was freely estimated from the data by anchoring the slopes at ages 8 and 14. Thus, the time-varying estimates represent deviations from the participants overall mean level and their potentially *non*-linear change across these ages.

Criterion 3—the final criterion pertaining to the stress exposure model—entails that the gender difference in depressive symptoms should be explained (i.e., mediated) by girls potentially becoming exposed to more stress than boys. We examined this criterion through mediation analyses using Sobel's test (Mplus does not enable bootstrapping with population weights).

Regarding the stress reactivity explanation, Criterion 4 states that SLEs and bullying victimization should predict depressive symptoms more strongly for girls than boys—and at the within-person level. Criterion 5 states that a stronger association for girls should be

specific to early adolescence (i.e., just before the gender difference in depression emerges). We estimated interaction terms between gender and SLEs and bullying victimization, respectively, in the ALT-SR models by following the procedures described by Mulder and Hamaker (2020). The interaction terms were added at ages 10 and 12. We inspected whether gender differences in within-person predictions between SLEs/bullying victimization and depressive symptoms were present from age 12 to age 14 (Criterion 4)—and only at this age and not from ages 10 to 12 (Criterion 5).

3. RESULTS

This chapter summarizes the main findings from Studies I-III. Detailed results, tables and figures are provided in the three studies/papers. Please note that all three studies were published or submitted with additional tables online. These can be inspected in the Appendix.

3.1.Study I

To shed light on the development of depression, we examined five types of stability from preschool into adolescence. MDD and dysthymia were investigated separately. (a) *Stability of form* was defined two-fold: i) structural stability and ii) frequency of symptoms. The CFA analyses indicated structural stability in both MDD and dysthymia across preschool into adolescence (see Study I, Table 2). A model with equal factor loadings of symptoms of dysthymia across ages 4 to 14 showed a better fit than a model in which factor loadings were set to be freely estimated. This result entails that the symptoms were of equal importance to the dysthymia construct across all ages. Evidence showed equal importance of the symptoms of the MDD construct across ages 4 to 12. However, the model fit was improved when factor loadings were freely estimated from ages 12 to 14—indicating increased importance of the symptoms to the MDD construct from ages 12 to 14.

Further on, the most frequent symptoms were irritability, change in weight/appetite, sleep disturbances, and diminished concentration (symptoms of MDD and dysthymia),

excessive guilt or feeling of worthlessness (symptoms of MDD), and low self-esteem (symptom of dysthymia) (see Study I, Table 1). The results are illustrated in Figure 2.

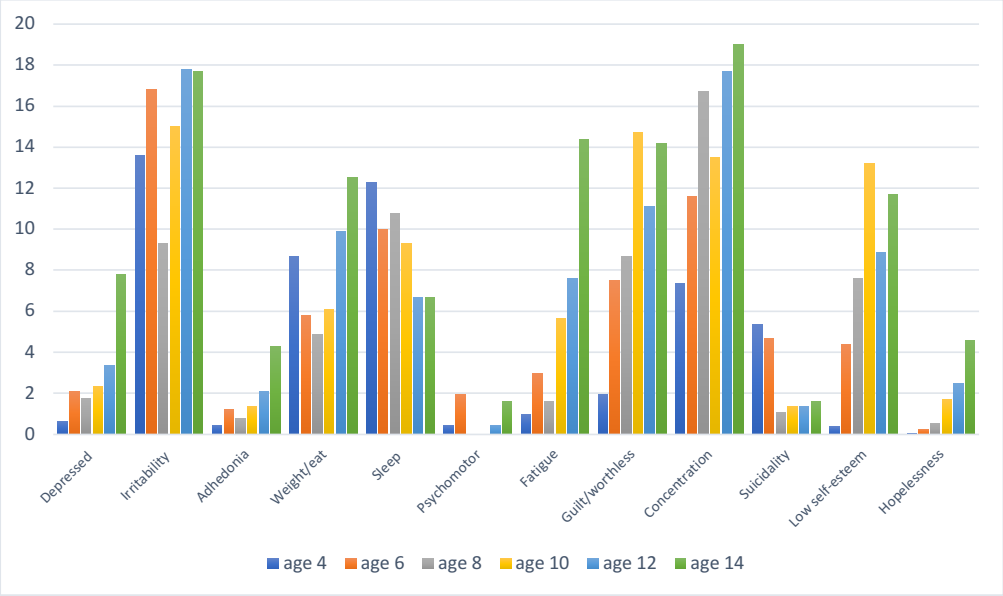


Figure 2: Frequency of symptoms of MDD and dysthymia.

Most symptoms evinced a linear and/or quadratic growth (see Study I, Table 1). The symptom of low self-esteem became more frequent towards age 10, a slight decrease in frequency at 12 and increased again at age 14. Sleep disturbances and suicidality *decreased* across ages 4 to 14. Notably, suicidality was highest in 4- to 6-year-olds, and thereafter relatively low-prevalent and stable (between 1.1 to 1.6%). We explored whether the higher prevalence of suicidality in the youngest ages could be explained by the age-adjustments we made in the PAPA interview (see section 1.1.). However, death themes in play were not common amongst 4- and 6-year-olds (1.7% and 0.5%, respectively).

Regarding (b) prevalence (i.e., group-level stability), we first examined increases in number of symptoms of MDD and dysthymia, thereafter we investigated MDD and dysthymia categorically (i.e., diagnoses). There was a linear and quadratic increase in

symptoms of MDD from ages 4 to 14 (see Study I, Table 3). There was neither a linear nor a curvilinear increase from ages 4 to 12—indicating that the increase in symptoms of MDD occurs between ages 12 and 14. There was a linear increase in symptoms of dysthymia from both age 4 to age 14 (see Study I, Table 3), indicating a steady increase from preschool until adolescence.

The prevalence of diagnosed (categorically measured) MDD was lower in the preschool age than adolescence (e.g., 0.11% at age 4 versus 2.68% at age 14). However, given the low prevalence of MDD at ages 8 and 10 we could not estimate a growth curve. The prevalence of diagnosed dysthymia was higher in adolescence (e.g., 5.08% in 14-year-olds) and showed a curvilinear increase across ages 4 to 14. We also found a linear increase also from age 4 to age 12. Thus, the curvilinearity was likely due to a sharp increase from 12 to 14.

Collectively, then, number of symptoms—and by all likelihood, diagnosis—of MDD increase from ages 12 to 14. Dysthymia (both the symptom-measure and diagnosis) increase from preschool until adolescence. Diagnosed dysthymia also shows an accelerated increase from ages 12 to 14.

Pearson's correlation analyses indicated modest to moderate (c) rank-order stability in both symptoms of MDD and dysthymia across ages 6 to 14 (see Appendix, Table S1, upper triangle). From age 4, rank-order stability was low, and lasted only until age 8. The stability was higher from ages 12 to 14 than from ages 10 to 12. This difference in stability did not emerge at earlier ages (i.e., not stronger from 10-12 than 8-10; from 8-10 than 6-8; or from 6-8 than 4-6).

The ICC analysis showed evidence for (d) absolute stability, and the results were comparable with the Pearson's *r* (see Appendix, Table S1, lower triangle). Whereas ICC is a direct expression of explained variance, Pearson *r* squared (r^2) indicates explained variance. For MDD, the average r^2 across age 4 to 14 was .09 and the average ICC was .22. For

dysthymia the average r^2 was .09, and the average ICC was .21. Thus, stability relative to oneself (absolute stability) was more than twice as high as stability relative to the group (rank-order stability). From ages 4 to 10, most children with depressive symptoms had fewer symptoms two years later, while about one in four children with no symptoms had acquired symptoms two years later (see Appendix, Tables S2-S6). Moreover, those with symptoms at age 12 were often inclined to remain at their level or acquire even more symptoms at age 14.

Finally, we investigated (e) stability of within-person changes with RI-CLPM. Across all ages, increased or reduced number of symptoms of MDD and dysthymia forecasted a corresponding increase or decrease two years later (see Study I, Table 4). Additionally, for both disorders, within-person increases or decreases at age 10 predicted similar changes *four* years later (age 14). For dysthymia, this four-year prediction also appeared for within-person changes at age 8 (thus predicting corresponding changes at age 12). For both MDD and dysthymia, the two-year predictions were stronger from ages 12 to 14 than from 10 to 12. At younger ages the predictions were of similar magnitude between adjacent assessment points.

3.2.Study II

In Study II, our overall aim was to inform on the predisposition and scar-explanations for the relationship between depression and personality. Specifically, we investigated a) whether increases or decreases in the Big Five personality traits predicted increases in depressive symptoms (MDD and dysthymia analyzed collectively) and b) whether increased depressive symptoms predicted changes in these traits, across ages 10 to 16. We used RI-CLPM and investigated the Big Five personality traits in five separate models. Each model included predictions in both directions: from personality traits to depressive symptoms and vice versa. We also included—and thus accounted for—potential time-varying effects of neuroticism (in the models examining the other four traits) and gender (in all five models). A secondary goal was to investigate developmental effects. If models with cross-lagged paths

set to be equal fitted the data as well as a freely estimated model, this would indicate no developmental effects. Finally, to compare our results with previous research, we investigated the RI-CLPM models without including neuroticism and gender as covariates, as well as running traditional CLPM analyses.

a) Within-person changes in personality traits forecasting increased depression would be in line with the predisposition model. Increased neuroticism predicted increases in depressive symptoms from ages 10 to 12 and 12 to 14—but not from ages 14 to 16. Changes in extraversion, conscientiousness, agreeableness, and openness did not predict future changes in depressive symptoms (see Study II, Table 3). We mostly found evidence for no developmental differences in the predictions across ages 10 to 16 (see Study II, Table 4). The one exception was that the paths from neuroticism on depressive symptoms were similar across the first two lags and differed from ages 14 to 16.

When the RI-CLPM models were rerun without controlling for gender and neuroticism, the results in which increased neuroticism predicted increases in depressive symptoms across ages 10 to 14 were replicated. Additionally, decreased agreeableness predicted increases in depressive symptoms across ages 10 to 16 (see Appendix, Table S22). The results on developmental differences were similar (see Appendix, Table S23).

Finally, CLPM-models replicated the original RI-CLPM findings on neuroticism. However, conscientiousness, agreeableness, and openness also predicted depression, and across all ages (see Appendix, Table S24). Thus, between-person results indicated that almost all traits pose vulnerabilities for depression, differing considerably from our original RI-CLPM results. Results on developmental effects, however, echoed the original results (see Appendix, Table S25).

b) Whether within-person increases in depressive symptoms forecasted within-person changes in personality traits would inform on the scar-model. Increased depressive symptoms

predicted increased neuroticism and decreased conscientiousness, from ages 10 to 12 and 12 to 14—but not from ages 14 to 16. Furthermore, an increased number of depressive symptoms predicted reduced extraversion at all lags (see Study II, Table 3). Results mostly indicated no developmental differences, except that the 14–16-age span for neuroticism and conscientiousness differed from the previous age-lags (see Study II, Table 4).

We reran the RI-CLPM without controlling for neuroticism and gender. The predictions echoed the original RI-CLPM results, with two exceptions. Increases in depressive symptoms predicted reduced agreeableness from ages 10 to 12, and conscientiousness only from ages 10 to 12 (not across ages 10 to 14 as the original results) (see Appendix, Table S22). The 14- to 16-age span for neuroticism still differed from previous ages. Moreover, the paths from depressive symptoms to extraversion, conscientiousness, and agreeableness varied across all ages, indicating developmental differences (see Appendix, Table S23).

When we reran the models with CLPM, the original results on neuroticism and conscientiousness (10 to 14)—extraversion (10 to 16)—as well as developmental effects were replicated. Additionally, increases in depressive symptoms predicted reduced agreeableness from ages 10 to 12 (see Appendix, Table S24) and this path differed from later age-lags (see Appendix, Table S25). That is, a potential scar mechanism from depression on reduced agreeableness only appeared at the between-person—and not the within-person—level.

3.3.Study III

The goal of Study III was to illuminate potential explanations for the emerging female preponderance in depression. Within the framework of a) stress exposure and b) stress reactivity explanations, we examined the potential role of SLEs and bullying victimization in the gender difference in depression. We proposed three criteria that need to be fulfilled to provide support for the stress exposure explanation and two criteria to indicate support for the

stress reactivity explanation. Children and adolescents were followed from age 8 until age 14, and within-person prospective associations were examined by ALT-SR models. Symptoms of MDD and dysthymia were analyzed collectively.

First and foremost, we found the expected female preponderance in depressive symptoms in early adolescence. The results showed low mean levels of depressive symptoms from ages 8 to 12 for both genders (between 1.0 and 1.5%), but with a sudden increase to 2.1% at age 14 for girls only (see Study III, Table 2, and Figure 2). Moreover, female gender correlated with depressive symptoms at age 14—and predicted an increase in depressive symptoms from age 12 to 14 (see Study III, Table 2)—but not earlier on. Thus, the female preponderance in depression emerged sometime between ages 12 and 14. Of note, at age 8, boys showed statistically significantly higher levels of depression than girls.

a) The stress exposure explanation was informed by three criteria. We first examined Criterion 1: whether girls become exposed to more stress than boys in the period before the gender difference in depression emerges (in our sample, the female preponderance emerged at age 14—so this would be the age-span 10 to 12). Both boys and girls experienced an increase in SLEs from ages 10 to 12, bullying victimization remained stable. There were no gender differences in exposure to either stressor (see Study III, Table 2); Criterion 1 was not fulfilled. Criterion 2 states that an increase in stress should predict an increase in depressive symptoms at the within-person level—at least among girls. We found that increases in SLEs and bullying victimization at age 12 predicted increases in depressive symptoms at age 14 in the ALT-SR analysis (see Study III, Figure 3). That is: Criterion 2 was supported. Regarding Criterion 3—whether the gender difference in depression at age 14 was mediated by increased stress—there was no mediation effect for neither SLEs (B=-0.01, *SE*=0.03, 95% CI [-0.08, 0.05]) nor bullying victimization (B=-0.04, *SE*=0.04, 95% CI [-0.12, 0.03]). In conclusion, criteria 1 and 3 pertaining to the stress exposure explanations were not fulfilled.

b) The stress reactivity explanation was investigated using two criteria. Criterion 4 states that stress should predict depression more strongly in girls than in boys. We did find that increases in both SLEs and bullying victimization at age 12 predicted increases in depressive symptoms at age 14—among girls but not among boys (see Study III, Figure 3). We also found statistically significant interaction effects between gender and age 12 SLEs and bullying victimization, respectively, on age 14 depression. The predictions were only statistically significant for girls. Finally, Criterion 5 states that the gender difference in the strength of predictions from stress on depression should not be present before the gender difference in depression emerges. In our sample—in which the female preponderance emerged at age 14—this would entail no gender-specific predictions from age 10 stress to age 12 depression. In Study III, Figure 3, the beta-values show no gender-specific predictive effects from age 10 to 12. Also, there were no interaction effects between gender and age 10 SLEs and bullying victimization, respectively, on age 12 depression. Thus, both criteria pertaining to the stress reactivity explanation were fulfilled.

4. DISCUSSION

4.1.Summary of findings

The overall aim of this thesis was threefold: 1) to investigate five types of stability in depression across preschool into adolescence; 2) to inform on the relation between adolescent depression and personality; and 3) to shed light on possible explanations for the female preponderance in adolescent depression. The thesis is based on longitudinal data from the Trondheim Early Secure Study (TESS) (Study I: ages 4-14, n = 1,042; Study II: ages 10-16; n = 817, Study III: ages 8-14; n = 748). Depressive symptoms were measured biennially with clinical interviews of symptoms of DSM-5-defined major depressive disorder (MDD) and dysthymia (analyzed separately in Study I and collectively in Study II and III). Predictions

were investigated by within-person methodology, accounting for time-invariant confounding effects.

In Study I, we investigated different types of stability in depression across ages 4 to 14. To inform on e) stability of form, we examined i) structural stability of the depression constructs and ii) which symptoms were most frequent at which ages. Further on, we investigated b) the prevalence of MDD and dysthymia (subthreshold symptoms and diagnoses) across preschool into adolescence. Finally, we investigated c) rank-order stability (i.e., staying at the same depression level compared to other children), d) absolute stability (i.e., staying at the same depression level over time), and e) stability of within-person changes (whether within-person changes in depressive symptoms predicted later corresponding changes). Within-person predictions would be in line with the hypothesis that depression leaves "depressive scars", which thereafter increase the risk of recurrence. Overall, we found evidence for all types of stability from preschool into adolescence—and increasingly so when entering adolescence:

- These depression constructs were coherent across ages. The most common symptoms were irritability and concentration difficulties.
- MDD and dysthymia became more prevalent in early adolescence.
- Rank-order stability was modest to moderate. Absolute stability was higher than rank-order. Within-person increases or decreases in depressive symptoms predicted later corresponding changes in depression.

In Study II, we investigated a) whether changes in the Big Five personality traits predicted increases in depressive symptoms, which would be in line with a predisposition explanation for the relationship between depression and personality. We also investigated b) if increases in depressive symptoms predicted changes in personality traits, which would be in line with a scar-explanation. Predictions were examined from ages 10 to 16, including

whether they differed across these ages. If so, this would indicate developmental differences. Results showed that:

- Increased neuroticism predicted increases in depressive symptoms—and increases in depressive symptoms predicted increased neuroticism—across ages 10 to 14.
- Increases in depressive symptoms predicted decreased conscientiousness from ages 10 to 14 and decreased extraversion across ages 10 to 16.
- With a few exceptions, the predictions did not differ across ages 10 to 16.

In Study III, we investigated whether girls became exposed to more SLEs and bullying victimization prior to the emerging female preponderance in depression and if a gender difference in stress exposure would mediate the gender difference in depressive symptoms. These findings would inform on the stress exposure explanation for the female preponderance in depression. Secondly, we examined whether increases in these stressors predicted increases in depressive symptoms more so for early adolescent girls than boys, which would be in line with the stress reactivity explanation. According to our findings:

- The female preponderance in depressive symptoms emerged at age 14.
- Girls did not become exposed to more of stressful live events (SLEs)—nor to bullying victimization—than boys.
- Increases in these stressors at age 12 predicted increases in depressive symptoms at age 14 more strongly for girls than boys. This interaction effect was not present at earlier ages.

4.2.Childhood depression poses a risk for adolescent depression

4.2.1. Stability of within-person change

In Study 1, we investigated stability of within-person change in depression from preschool until adolescence, using the participants as their own controls—accounting for time-invariant confounding effects. We found that increases or decreases in depression at one

age predicted corresponding later changes throughout preschool until adolescence. We investigated and found predictions across both two and four-year spans (i.e., increased depression at age 10 predicted increased depression at ages 12 and 14). The within-person predictions were stronger from ages 12 to 14 than previous age-lags. Of note, barring one within-person longitudinal study on early childhood (Wichstrøm et al., 2017), we were the first to investigate stability of within-person change, and our results should be replicated.

Even so, these findings are in line with the *scar hypothesis,* which proposes that earlier depression leads to changed characteristics of the child/adolescent—or their environment—which thereafter increases the likelihood of recurrence (Rohde et al., 1994). Although results from within-person analyses cannot be causally interpreted (Mund & Nestler, 2019), this approach is arguably more relevant for the scar hypothesis than between-person methodology.

Our results on the stability of within-person changes do not imply which "scar"-mechanisms might be involved in the recurrence of depression. Previous studies do allude to some possibilities. First, Coyne's interpersonal theory of depression (Joiner et al., 1999) proposes that depression leads to relational erosion. Depressed individuals may be characterized by negative interpersonal interaction styles—such as excessive reassurance seeking (Starr & Davila, 2008)—that elicit negative affect in others and thus more rejecting responses (Branje et al., 2010; Segrin & Dillard, 1992). Less available social support (Thompson et al., 2006) and loneliness (Dunn & Sicouri, 2022) might lead to depression. Relatedly, studies indicate that depression may lead to co-rumination, which has been shown to predict peer stress (Nolen-Hoeksema et al., 2008; Rose, 2021; Rose et al., 2017; Spyropoulou & Giovazolias, 2022). Moreover, Spyropoulou and Giovazolias (2022) found evidence to suggest that sadness rumination mediated the relationship between depressive symptoms and bullying victimization. Both peer stress (Hankin et al., 2015) and bullying victimization (Christina et al., 2021; Spyropoulou & Giovazolias, 2022) predict subsequent

depression. Notably, co-rumination and rumination entail to focusing on and discussing negative affect and problems (Nolen-Hoeksema et al., 2008; Rose, 2002). These tendencies might conceivably be especially pronounced in the early adolescence—marked by intense and prolonged negative affect (Rapee et al., 2019). Moreover, early adolescence is also a period of social reorientation and increased importance of peers (e.g., Harter, 2015; Somerville et al., 2013). It is thus possible that peer stress is even *more* predictive of depression in this age group than in younger children. This suggestion is in line with a study that indicates that peer stress could have a stronger impact on depressive symptoms in early puberty when compared to before puberty (Hankin et al., 2015). In conclusion, when treating a depressed adolescent, focusing on potential negative interpersonal interaction styles, including a tendency to co-ruminate and ruminate, as well as lack of social support might be important.

Another potential scar-mechanism relates to cognitive functioning. For example, one study on young adults with first onset MDD showed evidence of impaired inhibition and switching (Schmid & Hammar, 2013). These impairments persisted a year later even though symptoms of depression had receded, and those participants with the most cognitive impairment in the acute phase of their MDD episode had an increased likelihood of relapsing within a year (Schmid & Hammar, 2013). Another study on adults did not find persistent impartment in these abilities after five years (Ronold et al., 2020)—thus, it is possible that the changes are temporary—at least in adults. Halse et al. (2022) found that in the TESS-sample, increased levels of depression predicted increased cognitive inflexibility in young children two years later. Executive functioning is strongly associated with the functioning of the prefrontal cortex (PFC)—which develops until early adulthood (e.g., Somerville et al., 2013). Therefore, it could be more likely that depression leaves scars in the form of reduced executive functioning in adolescents than adults. This possibility should, however, be investigated.

4.2.2. Rank-order and absolute stability

Most prior research on the risk of recurrence has investigated whether depression increases the risk of a new depressive episode when compared to those who have not been depressed (i.e., rank-order stability). These studies are mainly conducted on adult samples (Richards, 2011) or they have investigated rank-order stability from adolescence into adulthood (e.g., Johnson et al., 2018). In Study I, we were the first to investigate rank-order stability in symptoms of DSM-defined MDD and dysthymia from age 4 until 14—thus including the age when depression typically first increases in prevalence (Ford et al., 2003; Merikangas et al., 2010a). We found modest to moderate rank-order stability in depressive symptoms from age 6 until age 14, and low from age 4 until age 8. Rank-order stability increased from age 12 to 14 when compared to previous ages. Thus, already from early childhood on, there is a risk of keeping one's "depression position" in the group—and even more so when entering early adolescence. Our study expands on previous studies showing rank-order stability from preschool until age 12 in a collective measure of symptoms of MDD, dysthymia, and depression not otherwise specified (NOS) (Finsaas et al., 2018), and from 11 to 18 in depressive symptoms measured with a questionnaire capturing only some DSM-defined symptoms (Mason et al., 2017). The one study investigating categorically defined depressive disorders in the general population did not find evidence for rank-order stability from the age-groups 9-12 to 13-18 (Copeland et al., 2013a). This last diverging finding may stem from low statistical power of the study.

Furthermore, in Study I, we investigated absolute stability in depression (i.e., the probability of staying at ones' *own* depression level over time). We found evidence that absolute stability was stronger than the rank-order stability. That is, the use of correlational methods (i.e., Pearson *r* or OR) when the aim is in fact to investigate absolute stability, correlational methods will underestimate the stability. We also inspected how many youths

with symptoms at one age had equally as many symptoms at later ages (see Appendix, S2-S6). When entering adolescence, more adolescents acquired new symptoms than children at previous ages did. For example, of those individuals with 3 symptoms at age 6, only 2% had increased to 4 symptoms at age 8, while 34% had reduced to 1 symptom. However, of those with 3 symptoms at age 12, 38% had increased to 4 symptoms at age 14, while 26% had reduced to 1 symptom. In other words: 12-year-olds had more often acquired more depressive symptoms and less often outgrown their depressive symptoms by age 14.

Collectively, we found evidence for modest to moderate rank-order stability, stronger absolute stability, and that adolescents "depression position" relative to the group and their own level fortifies in early adolescence.

4.3.Prevalence and stability of form across preschool to adolescence

4.3.1. Prevalence

The prevalence rates in Study I are mostly compatible with previous research showing that MDD and dysthymia diagnoses are very low prevalent in the preschool age (Lavigne et al., 2009; Wichstrøm et al., 2012), fairly low prevalent in middle childhood (Costello et al., 1996; Ford et al., 2003; Merikangas et al., 2010a), and higher in early adolescence (Avenevoli et al., 2015; Gau et al., 2005; Merikangas et al., 2010a). However, in our data, diagnosed dysthymia was more common than MDD—especially in adolescence (e.g., at age 12: dysthymia 2.02% versus MDD 0.71%; and at age 14: dysthymia 5.08% versus MDD 2.68%). In contrast, previous research shows a *lower* prevalence of dysthymia (0.2%-1.1%) (Gau et al., 2005; Merikangas et al., 2010a) compared to MDD (3.8%-5.5%) (Avenevoli et al., 2015; Merikangas et al., 2010a) in comparable ages. Notably, the study with the lowest prevalence of dysthymia (Gau et al., 2005) was conducted in Taiwan, and the other studies were conducted in the United States. Thus, it is possible that this research is not comparable to ours. In conclusion, our findings indicate that dysthymia may be more common than MDD in

adolescence, at least in a northern European context. According to Schramm (2020), dysthymia is more easily overlooked because of its milder presentation than other forms of chronic depression. Even so, it is associated with functional impairment and serious outcomes such as increased risk for suicide (reviewed by; Schramm et al., 2020) and later MDD (Klein et al., 2006). Our finding underscores the importance of identifying dysthymia in children and adolescents—in research as well as clinical practice and preventative efforts.

Beyond inspecting the prevalence rates of dysthymia and MDD, we estimated growth-curves (of both mean level symptoms and diagnosis) to investigate more precisely *when* the increases in these disorders occur. Our results indicate a stable prevalence of on subthreshold symptoms of MDD across earlier childhood, with an increase from ages 12 to 14. Even though growth curves could not be estimated for diagnosed MDD given the low prevalence at ages 8 and 10, the prevalence was higher at ages 12 (0.71%) and 14 (2.68%) than at earlier ages. Our findings are in line with research showing an increase in diagnosed MDD in early adolescence (Ford et al., 2003; Merikangas et al., 2010a). Notably, Costello et al. (1996) did *not* find any increase across ages 9 to 13. Thus, it is possible that the increase in MDD first occurs between ages 13 and 14. Regarding dysthymia, our study is the first to investigate and demonstrate a steady increase already from the preschool age and into adolescence—both in subthreshold symptoms and diagnosis. Moreover, diagnosed dysthymia showed an even *escalating* increase from ages 12 to 14—in contrast to studies showing no change in dysthymia when entering adolescence (Costello et al., 1996; Merikangas et al., 2010a). Because the prevalence in our study was 2-5 higher than in these prior studies, we may have had more statistical power to detect these differences.

4.3.2. Structural stability of the depression construct

A second aim of Study I was to illuminate whether the depression construct appears the same across preschool into adolescence. This was first investigated as *structural stability*

(i.e., by partial measurement invariance). Previous studies on structural stability have provided inconsistent results and have not examined the DSM constructs of MDD and dysthymia per se (Lahey et al., 2004; Weiss & Garber, 2003). We found evidence in support of structural stability of both MDD and dysthymia across preschool into adolescence. The factor loadings of the nine symptoms of MDD did not change across ages 4 to 12 and the seven symptoms of dysthymia did not change across ages 4 to 14. These findings indicate that individual symptoms of depression are equally important to the overall MDD and dysthymia depression constructs across these age-groups. For example, irritability may likely be as important in indicating depression at age 6 as at age 12. Moreover, although factor loadings in symptoms of MDD did not change across earlier childhood, we did find increased loadings from age 12 to 14. This suggests that the MDD construct becomes even *more* coherent from adolescence on. That is: when entering adolescence, a single symptom is more likely to be accompanied by more symptoms—perhaps a full-blown depression diagnosis. An implication is that when a youth evinces one depressive symptom, referral for clinical assessment is perhaps *especially* warranted in early adolescence.

Our findings indicating structural stability from preschool into adolescence have three other important implications. First, they indicate that our results on stability of within-person changes; rank-order; and absolute stability, represent stability in the same underlying construct. Second, they support the notion that the DSM-5 system may be used to diagnose MDD and dysthymia in both children and adolescents. Third, and relatedly, our results provide evidence in support of the age-adjustments of the DSM-5 (e.g., that the core symptom might be persistent irritability), in that irritability loads on the depression construct and to an equal degree across preschool into adolescence. It is, however, important to note that we did not directly address the *developmental appropriateness* of these criteria. It is possible that other symptoms not captured in the DSM-5 system may be relevant for depression in children

and adolescents. Nor did we investigate whether some of the symptoms might be even more important to *other* constructs within the DSM-system, e.g., irritability for Disruptive Mood Dysregulation Disorder (DMSS) (American Psychiatric Association, 2013). Finally, symptoms might be expressed in different ways in children versus adolescents, which is not described in the DSM-5.

4.3.3. The most common symptoms

A second aspect of how depression presents across childhood into adolescence is how typical symptoms are in a certain age group: *frequency of symptoms*. Our results revealed that some of the symptoms were the most common in most of the age-groups. Moreover, almost all symptoms became more frequent with age—shown by linear and quadratic increases.

Irritability was the most common symptom across all ages. As previously described, the age-adjustment of irritability as a core-symptom for MDD and dysthymia in the DSM-5, is not included in the ICD-system. Moreover, across ages 4 to 12, depressed mood and anhedonia were fairly infrequent. These are the other two core-symptoms in the DSM-5 and the only core symptoms in the ICD-11. Collectively, these findings have an important implication: clinicians who use the ICD, such as in the current context (Norway), should be mindful of irritability as a potential symptom of depression in children and adolescents. If not, they run the risk of missing depressive episodes. The second most common symptom in our sample was concentration difficulties. Concentration difficulties, or inattention, are one of the strongest markers of attention-deficit/hyperactivity disorder (ADHD), especially in adolescence when the symptoms of impulsivity/hyperactivity typically decrease (Eng et al., 2023; Fraticelli et al., 2022). There is also a considerable comorbidity with depressive disorders (ADHD three times more common in people with depressive disorders; Sandstrom et al., 2021). Thus, when clinicians assess or treat children or adolescents with ADHD, a

worsening of concentration difficulties in periods of persistent low mood/irritability and/or anhedonia—could indicate depression, or comorbidity with depression.

Unexpectedly, suicidality (ideation, behaviour, attempts) was highest in 4- to 6-year-olds and thereafter became less prevalent. We explored the possibility that the higher prevalence in these ages stems from the age-adjustments for this symptom done in the PAPA-interview. However, death themes in play were very rare in our sample (see section 3.1.). Another possible explanation is related to preschoolers asking many questions of their parents, including about death (e.g., Chouinard et al., 2007). This normative occupation with death may leave some parents worried and more likely to respond affirmatively to questions regarding thoughts about death. However, the factor loadings for suicidality did not change across ages. Therefore, it is possible that we did capture suicidality as a symptom affecting 5% of the youngest children. The low prevalence of suicidality in adolescents was arguably even more notable. Other studies rather show an *increase* in suicidal ideation from late childhood into adolescence (Zhu et al., 2019), and that ideation, plans, and attempts increase towards mid-adolescence for girls and towards late adolescence for boys (e.g., Boeninger et al., 2010). One possible explanation for the low prevalence of suicidality in our sample may be related to the use of clinical interviews. The adolescent is likely fully aware of the seriousness of the topic and might not want to answer openly to avoid worrying their surroundings. For the same reason, the adolescents may not have told their parents about these symptoms, and parent reports may not necessarily add information about suicidality. Moreover, adolescence is a period of social reorientation from parents to peers (e.g., Harter, 2015), and parents might therefore be less privy to the psychological (e.g., unobservable) symptoms of depression in general (Weiss & Garber, 2003)—including suicidal thoughts. Previous work has mainly used self-reported questionnaires. This is possibly a better way of capturing this particular symptom—all things considered.

Another surprising finding in our data is the *decrease* in sleep-difficulties. A previous meta-analysis found that the depressive symptom of hypersomnia was more common in adolescents than children, and insomnia was as frequent (Weiss & Garber, 2003). Moreover, many adolescents struggle with sleep-debt (Rapee et al., 2019). However, sleep-difficulties as a depressive symptom must occur as part of a depressive episode and is therefore distinct from both a clinical hypersomnia/insomnia diagnosis as well as normative sleep-debt issues.

Finally, low self-esteem was most frequent from ages 10 to 14. The normative decrease in global self-esteem from early adolescence and on has been considered a product of an increased ability for perspective taking (Enright et al., 1980) and concern with how one is being evaluated by others (Harter, 2015; Westenberg et al., 2004). These developmental changes might also increase the likelihood of low self-esteem as part of a depressive episode occurring in adolescence rather than earlier childhood.

Collectively, these findings indicate that DSM-5-defined MDD and dysthymia mostly appear the same across preschool into adolescence. However, structural stability and most symptoms did increase when entering adolescence—paralleling an increase in stability of within-person change, rank-order and absolute stability, and prevalence. Collectively, the results from Study I underscore the importance of implementing preventative and treatment efforts in early adolescence, and of gaining knowledge on which factors contribute to adolescent depression. This was the focus of Studies II (on personality) and III (on the female preponderance).

4.4.Adolescent depression and personality

In Study II, our aim was to inform on two different explanations for the relation between depression and personality: whether personality might entail vulnerability for depression (a predisposition explanation) and whether depression may impact personality development (a scar explanation). Across ages 10 to 16, we investigated longitudinal within-

person predictions between depressive symptoms (symptoms of MDD and dysthymia analyzed collectively) and the Big Five personality traits: neuroticism, extraversion, conscientiousness, agreeableness, and openness.

4.4.1. Reciprocal relations between adolescent depression and neuroticism

First of all, we identified that increases in neuroticism predicted increases in depressive symptoms—and vice versa—across ages 10 to 14. These findings are in line with both predisposition and scar-explanations: high levels of neuroticism may pose a vulnerability for developing depression in early adolescence, and depression may impact personality development in the form of increased neuroticism. Our within-person results solidify previous studies showing these reciprocal predictions at a between-person level (e.g., Goldstein et al., 2020; Klimstra et al., 2010; Zhang et al., 2020), and expand on prior studies by—for the first time—investigating a sample as young as age 10. Notably, in Study I, we found that increases in depressive symptoms at age 10 predicted increases in depressive symptoms at both age 12 and 14. These findings from Study I and II combined tentatively indicate a scar-process in which increased depression at age 10 leads to increased neuroticism at age 12 (scar) that thereafter poses as a vulnerability factor for depression at age 14.

Even though we did not investigate possible mechanisms explaining how neuroticism might pose as a vulnerability factor for depressive symptoms—and how depressive symptoms might leave scars in the form of increased neuroticism—previous research alludes to several possibilities. For example, adolescents scoring high on neuroticism may respond to stress in maladaptive ways, which increases their risk of depression. Indeed, studies have shown that early adolescents scoring high on neuroticism are more prone to self-blame (Liu et al., 2020) and emotional suppression (Yoon et al., 2013)—all strategies increasing the risk for depression. They may also *perceive* more stress, for example, school stress (Hoferichter et al., 2014). In fact, Tian et al. (2019) found that perceived school stress mediated the relationship

between neuroticism and depression. Finally, neuroticism has been associated with low self-esteem (Zeigler-Hill et al., 2015)—possibly because of its similar emotional underpinnings (DeNeve & Cooper, 1998), and low self-esteem is also a risk-factor for adolescent depression (Masselink et al., 2018).

As regards possible scar-mechanisms from depressive symptoms to neuroticism, rumination may be involved. For example, Ronold et al. (2020) found that depression predicted higher levels of rumination, and this change held up to five years after depression had subsided. Moreover, rumination has consequently been associated with neuroticism (Liu et al., 2020; Vidal-Arenas et al., 2022), and may thus explain how increased depression might lead to increased neuroticism. Another possibility is that depression leaves "self-critical" cognitive scars. Depression may alter the way individuals process self-relevant information in that they are more likely to attend to, encode, and retrieve negative information about themselves (Mu et al., 2019). Self-criticism has been proposed to be organized under (at the facet level) the higher-order trait of neuroticism (Clara et al., 2003).

Of note, in Study II, we also found that the reciprocal predictions between depressive symptoms and neuroticism were no longer statistically significant in the 14 to 16-age span. This will be discussed below (see section 4.4.4.).

4.4.2. Scars on extraversion and conscientiousness

Beyond that, increases in depressive symptoms predicted increases in neuroticism. In Study II, we also found increases in depressive symptoms to predict decreases in extraversion across ages 10 to 16—and decreases in conscientiousness across ages 10 to 14. The prediction on decreased conscientiousness was non-significant from ages 14 to 16, which I will elaborate on below (section 4.4.4.). Nevertheless, our study extends the one previous between-person study showing that depression predicted lower levels of these two traits from age 12 and forward (Klimstra et al., 2010). By using within-person methodology, our study solidifies

support for these scar-explanations, and extends the findings by indicating they may apply already from age 10.

Potential scar-mechanisms explaining how depressive symptoms might lead to reduced extraversion involve how depression might have a negative impact on the child's or adolescent's relationships (e.g., social erosion). First, depressive symptoms such as fatigue, low self-esteem, guilt, and loss of interest may directly influence how often the child socializes. Furthermore, this social withdrawal may lead to lower friendship quality and, thus, less interest and drive to seek contact with these friends after the depressive episode or symptoms have subsided. Overall, this may entail lower scores on extraversion. Secondly, some of the depressive symptoms involve a negative view of oneself as well as hopelessness. These may hinder the adolescent's ability to perceive social support that is actually present (Thompson et al., 2006), interpret social events in a negative and self-critical manner (Werner et al., 2019), and ultimately lead to further social withdrawal and lower scores on extraversion.

There are also several scar-mechanisms that may explain our finding that increased depressive symptoms predicted reduced conscientiousness. One possibility involves executive functioning. As previously noted, some evidence suggests that MDD might lead to cognitive inflexibility in young adults (Schmid & Hammar, 2013), and in very young children (Halse et al., 2022). Correspondingly, cognitive inflexibility has been associated with reduced conscientiousness (Fleming et al., 2016). However, to illuminate whether depression might leave "executive scars" that also decrease conscientiousness, we need prospective studies on cognitive inflexibility and other types of executive functioning, and in the adolescent period specifically.

In Study II, we did not find that reduced extraversion and conscientiousness predicted increases in depressive symptoms—indicating that these traits do not pose as vulnerability

factors for depression. Our findings solidify previous between-person results (Calvete et al., 2016; Goldstein et al., 2018; Yang et al., 2008), with one exception. Klimstra et al. (2010) found that lower levels of extraversion and conscientiousness predicted higher levels of depression. Even so, in contrast to Goldstein et al. (2018), Yang et al. (2008) and our study, Klimstra et al. (2010) did not account for the effects of neuroticism. As part of Study II, we both investigated predictions without accounting for neuroticism—as well as by using between-person methodology. In these models, extraversion and conscientiousness still did not predict depressive symptoms. Thus, the two main elements of the Klimstra et al. (2010) study that differed from our study did not seem to explain the diverging results. One final difference was that Klimstra et al. (2010) used self-reported questionnaires for both depressive symptoms and personality traits, possibly inflating the predictions from these traits on depression—but this is uncertain. Either way: our results fall in line with the majority of between-person studies indicating that decreases in extraversion and conscientiousness do *not* predispose for depression in adolescence.

4.4.3. Agreeableness and openness

Finally, for the personality traits agreeableness and openness, no statistically significant paths were identified. These null findings are in line with two other studies that have shown that reduced agreeableness and openness did not predict higher levels of depression (Goldstein et al., 2018; Klimstra et al., 2010). It has been suggested that individuals low on openness might show less adaptability when faced with difficult situations, and thus become more vulnerable to depression (Khoo & Simms, 2018). Prior research has also shown that low levels of agreeableness are associated with lower perceived social support (Barańczuk, 2019), a likely risk factor for depression (Rueger et al., 2016). However, our findings from Study II do not support the notion that these traits pose as neither vulnerability nor scars, at least in adolescents.

Our findings converge with the one previous study relevant for the *scar*-explanation showing that higher levels of depressive symptoms did not predict openness (Klimstra et al., 2010). However, in contrast to our null finding on agreeableness, Klimstra et al. (2010) *did* find that depressive symptoms predicted lower levels of agreeableness. Again, we reran the models without accounting for neuroticism as well as by using between-person methodology. In our sample, depressive symptoms still did not to predict agreeableness across ages 12 to 16—the age-groups overlapping with the Klimstra et al. (2010) study (baseline age 12 and onwards). As discussed regarding the diverging findings on extraversion and conscientiousness, it is unclear whether that common-methods effects in the Klimstra et al. (2010) study might have inflated their result.

4.4.4. Developmental differences

A secondary goal in Study II was to investigate whether predictions between depressive symptoms and the Big Five personality traits differed across ages 10 to 16. Our results mostly indicate no developmental differences. There were, however, two exceptions. Reciprocal predictions between depressive symptoms and neuroticism—and predictions from depressive symptoms on reduced conscientiousness—were no longer statistically significant from ages 14 to 16 and differed from previous age-lags.

It is possible that the reciprocal relation between depression and neuroticism—and the potential scar-effect on conscientiousness—only exists in the earliest adolescent years. However, with one exception (Williams et al., 2021), previous studies have shown reciprocal predictions between depression and neuroticism—and predictions from depression on conscientiousness—after age 14 (e.g., Goldstein et al., 2020; Klimstra et al., 2010; Yang et al., 2008). These studies have all used between-person methodology, but when we reran our models with between-person methods our original RI-CLPM null-findings from age 14 to 16 remained. Thus, the diverging results were not explained by the use of between-person versus

within-person methodologies. Therefore, why these effects should disappear after age 14 is unclear.

One possibility is that effects in the 14- to 16-age range would have been captured with shorter measuring time spans. Variants of the predisposition and scar explanations propose that the personality traits pose a vulnerability for increased depression, but only temporarily (a state-model; Ormel et al., 2013); and that depression may impact personality traits, but also only temporarily (a complication model; Ormel et al., 2013). Given that CAPA and Kiddie-SADS probe for symptoms occurring in the prior three months, and that we measured personality and depression every second year, effects might have subsided from age 14 until the three months prior to the age 16 assessment. Future research should explore whether predictions appear at shorter measurement intervals after age 14 and, if so—why effects appear to be of shorter duration at this age and not in the early adolescent years.

A possible explanation for the non-significant prediction from neuroticism on depression specifically relates to the lower prevalence of depressive symptoms at age 16 than at ages 14, 12 and 10 in our sample (see Study II, Table 2)—while previous research has consequently demonstrated that depression increases towards middle adolescence (e.g., Merikangas et al., 2010b). From ages 14 to 16, we changed the clinical interview used to measure depressive symptoms from the CAPA to the Kiddie-SADS. The Kiddie-SADS have somewhat stricter criteria in that a depressive symptom must have been present most of the day/at least 50% of the day—while in the CAPA the symptom must have been present at least *one hour* of the day (for a further discussion, see section 4.6.3.1.). The stricter Kiddie-SADS criteria may explain the apparent decrease in depressive symptoms from age 14 to 16, which may ultimately increase the risk of Type 2 errors (false null findings) in predictions from personality traits on depressive symptoms in this age range.

4.5.Adolescent depression and gender

Paralleling the striking increase in depressive symptoms in early adolescence, a female preponderance in depression emerges. The aim of Study III was to inform on the stress exposure and stress reactivity explanations for this gender difference. The stressors we examined were stressful life events (SLEs) and bullying victimization. We proposed and investigated specific criteria that need to be fulfilled to provide tentative support for the exposure (three criteria) and reactivity (two criteria) explanations. Predictions were examined at a within-person level across ages 8 to 14—to capture the period before the female preponderance emerges.

First of all, we found the expected gender difference in depressive symptoms (MDD and dysthymia analyzed collectively) at age 14. This corroborates a meta-analysis indicating that the female preponderance in depressive symptoms and MDD diagnosis emerges by age 12 at the earliest and thereafter increases towards middle adolescence (Salk et al., 2017).

4.5.1. Stronger depressive reactions in girls

Furthermore, in Study III, both criteria for the stress reactivity explanation were fulfilled. We found that increases in SLEs and bullying victimization at age 12 predicted increases in depressive symptoms at age 14 for girls but not boys. We also did interaction analyses and found that these predictions were stronger for girls than boys. Collectively, our findings are in line with a reactivity explanation for the emerging gender difference in depression: when early adolescent girls experience stress, they react more strongly than boys in the form of more depression.

First and foremost, by including a younger age-sample and using within-person methodology, these results solidify results from a previous between-person study indicating a stronger prediction from SLEs on depression in a sample of mean age 13 at first assessment (Ge et al., 1994). Second, our results extend mixed findings on whether bullying victimization

predicts depression more strongly in girls or only in girls (studies with baseline measurements at age 13) (Bond et al., 2001; Christina et al., 2021; Lepore & Kliewer, 2019). Our findings also add to a broader research base, consequently showing stronger between-person predictions or correlations between stress and depression for girls than boys in early adolescence, paralleling the age when the female preponderance in depression also emerges (typically around ages 12-13). Examples include stronger predictions from daily hassles and episodic stress in girls than boys (Bastin et al., 2015; Hankin et al., 2007; Shih et al., 2006), and online and offline sexual harassment that has been shown to correlate with depression in girls but not boys (Ståhl & Dennhag, 2021). In future studies, these stressors may be investigated with longitudinal within-person designs with a younger baseline age.

Our study did not examine the underlying mechanisms for the increased stress reactivity in girls. However, several possibilities may be drawn based on previous research. First, the pubertal transition has been shown to predict psychopathology—over and above chronological age (Mendle et al., 2020; Pfeifer & Allen, 2021; Stumper & Alloy, 2021). The majority of studies find that entering puberty is associated with increased risk for depression in girls but not boys (reviewed by Stumper & Alloy, 2021). Moreover, studies have found that SLEs (Ge et al., 2001) and bullying victimization (Compian et al., 2009) are more strongly associated with depression among adolescent girls that are more advanced in their pubertal development (i.e., “early timers”)—further implying the role of puberty in reactivity and depression in early adolescent girls. Some have suggested that the neural and physiological changes that come with puberty make adolescents more reactive to stress (Dahl & Gunnar, 2009; Spear, 2009). However, such an explanation has not been tested (Stumper & Alloy, 2021). Whether and how pubertal changes specific to girls might explain why they become more reactive to stress than boys, is also unclear. Several hypotheses might be suggested.

First, pubertal hormones might be involved. Prior research suggests a stronger link between pubertal hormones and depression in girls than boys (reviewed by Stumper & Alloy, 2021). Even though one recent study on girls only found that testosterone, not estradiol, predicted depression (Copeland et al., 2019), most research indicates the role of estradiol (Stumper & Alloy, 2021). Both stress (Doom & Gunnar, 2013) and pubertal hormones (e.g., Roberts & Lopez-Duran, 2019) influence the functioning of the body's main stress-system: the hypothalamus-pituitary-adrenal (HPA) axis. When activated by stressors, the HPA-axis initiates a cascade of behavioral and hormonal responses (e.g., increases in cortisol). In the short-term, these responses allow the body to respond to challenges in an effective and adaptive manner (McEwen, 1998). Based on a literature review, Roberts and Lopez-Duran (2019) concluded that from puberty and on, boys produce more testosterone and dehydroepiandrosterone (DHEA) than girls do. These hormones inhibit HPA-axis responses to stress and allow for adaptive regulation. Girls, on the other hand, produce more estradiol, which has excitatory effects on the HPA-axis (stress reactivity) (Roberts & Lopez-Duran, 2019). This might contribute to explaining the findings that the HPA-axis is more readily activated in adolescent girls than boys (Oldehinkel & Bouma, 2011). Over time, frequent or chronic activation may lead to hyperactive reactivity or hypo-reactivity to stress (e.g., blunted stress-responses) (Guilliams & Edwards, 2010). What remains elusive is how a more frequent activation of the HPA-axis in girls—partly driven by pubertal hormones—might lead to more depression (Oldehinkel & Bouma, 2011; Roberts & Lopez-Duran, 2019). Research has shown that adolescent girls evince lower cortisol levels (Bouma et al., 2009)—and more negative affect (Bailen et al., 2019)—than boys when exposed to social stress—indicating the role of hypo-reactivity. Oldehinkel and Bouma (2011) suggest that, from a theoretical point of view, a blunted cortisol response may leave the body with some specific depressive symptoms (e.g.,

more fatigue and less energy), which thereafter increase the risk of developing other additional depressive symptoms.

A second way puberty might be involved in increased stress reactivity in girls is that puberty increases the likelihood of *other* risk-factors for depression. For example, research indicates that secondary pubertal characteristics, including weight gain, increase body dissatisfaction in girls, particularly when entering adolescence (reviewed by; Hyde & Mezulis, 2020). Moreover, girls are more self-conscious (Rankin et al., 2004) and score higher on body-objectification (Grabe et al., 2007). Hyde and Mezulis (2020) suggest that different vulnerabilities—such as bodily pubertal changes combined with increased body objectification and body dissatisfaction—collectively represent an increased "depressogenic vulnerability" in adolescent girls, which will increase the overall likelihood of depression when faced with stress. In this context, the social media may further exasperate body dissatisfaction in girls. In a study based on the same sample as herein (TESS), it was reported that exposure to other-oriented content on social media predicted increased body dissatisfaction in girls but not boys (Steinsbekk et al., 2021). It has also been demonstrated that adolescent girls show an increased level of self-criticism in general, shame, and fear of self-compassion, and that relational stress mediates the prediction from these factors on self-harm (Xavier et al., 2016). Arguably, it is possible that girls—because of their higher overall self-criticism—are more vulnerable when exposed to stressors such as SLEs and bullying victimization. Research shows that self-critical individuals may be less likely to perceive available social support (Satterwhite & Luchner, 2016). Thus, highly self-critical adolescents—more often girls than boys—may be more vulnerable to a self-defeating attitude when facing stress—and less likely to perceive social support (Rueger et al., 2016): both of which increase the risk for depression.

Another potential mechanism concerning stress reactivity in girls relates to sleep. Difficulties with sleep patterns are more common in girls from early adolescence and on (Salk et al., 2017)—and female pubertal hormones are potentially involved in this gender difference (Morssinkhof et al., 2020). Stress likely deteriorates sleep-quality (Li et al., 2019), and a recent study showed that shorter sleep-duration predicted depressive symptoms in girls but not boys during adolescence (Mathew et al., 2019). Collectively—the potentially stronger reactions to SLEs and bullying victimization in girls might be mediated by lower sleep-quality.

Finally, gendered stress reactivity may be related to girls' use of rumination when regulating stress. Evidence suggests that girls ruminate more than boys—especially in early adolescence when compared to late childhood (Hampel & Petermann, 2005). Another study found that in early adolescence, rumination predicted depression in girls but not in boys (Krause et al., 2018). Moreover, early adolescent girls likely enter into more *co*-rumination with friends (Rose et al., 2017; Stone et al., 2011), which increases depressive symptoms to a stronger degree in girls than boys—possibly mediated through a further increase in interpersonal stress (Rose et al., 2017). Another study also found that the potential trade-offs with co-rumination—increased friendship quality but also increases in depressive symptoms—were particularly pronounced for girls Tilton-Weaver and Rose (2023).

4.5.2. Equal stress exposure

The results from Study III did *not* indicate support for the stress exposure explanation for the gender difference in depression. Although SLEs became more common from ages 10 to 12 (i.e., just prior to the emerging gender difference in depression), this increase was equal for girls and boys. Bullying victimization did *not* increase—for neither girls nor boys. Thus, the first and essential criterion pertaining to the stress exposure explanation was not fulfilled. The evidence for a gender difference in SLEs in early adolescence is mixed. Some studies

have found no female preponderance in SLEs in early adolescence (Jenness et al., 2019; Sund et al., 2003) while an older meta-analysis showed a stronger increase in adolescent girls than boys (Davis et al., 1999). As regards bullying victimization overall, results are also mixed, with one study showing no gender difference (Sweeting et al., 2006), while a Norwegian study showed an increase among girls only (Wendelborg, 2020).

Even though we did not find a gender difference in stress exposure in Study III, when entering early adolescence, girls might become more exposed than boys to *other* stressors—potentially contributing to the female preponderance in depression. For example, research has shown that early adolescent girls are more exposed to peer stress in general (Rose & Rudolph, 2006), school stress (Klinger et al., 2015), and daily interpersonal hassles (Hankin et al., 2007)—and all these stressors predict depression in girls. However, all these studies have baseline ages that parallel the emerging female preponderance in depression. Thus, these stressors may be investigated with a younger age-span and preferably by using within-person methodology.

4.5.3. Some perspectives on gender

4.5.3.1.Artifact explanations

As noted in the introduction, some have suggested that the female preponderance in depression may be exaggerated or is even an artifact. The *artifact hypothesis* proposes that the society is more accepting of demonstrations of vulnerability in girls than boys, and that girls therefore express their depression more readily (Parker & Brotchie, 2010). According to Motro and Ellis (2017) the male role discourages crying, whereas the female role might even encourage it. As a result, boys or young men might be less likely to report feelings of sadness, restrain themselves from crying, and/or not inform their surroundings when they are struggling. This may extend to when adolescent boys participate in research. Thus, the female

preponderance in adolescent depression may partly or completely be a result of gender-biased responding.

There is some research indicating such gender biases. Romans, Tyas, Cohen, and Silverstone (2007) found that women with clinical depression were more likely to report 4 out of 26 possible symptoms than men with clinical depression: “increased appetite”, “often in tears”, “loss of interest” and “thoughts of death”. A meta-analysis found that depressed women more often present with depressed mood, changes in weight or appetite, and sleep-disturbances (Cavanagh et al., 2017). Of these symptoms, bias in *often in tears* and *depressed mood* are in accordance with the artifact hypothesis.

Moreover, some information may be drawn from research showing gender difference in help-seeking behavior. A comprehensive report of 3000 adolescents showed that those with emotional problems sought informal (family, friends) rather than formal help (mental health professional). Help-seeking in general was somewhat more common in girls (97%) than boys (89%) (Gray & Daraganova, 2017). In line with previous findings (Rickwood et al., 2005), girls more often sought help from friends, while there was an equal gender balance in help-seeking from family and teachers (Gray & Daraganova, 2017). Research on victims of bullying is mixed: most show that girls more often seek help (Blomqvist et al., 2020; Cortes & Kochenderfer-Ladd, 2014) while others have found that boys do (Shaw et al., 2019). Collectively, there is some evidence to suggest that women more often present with depressive symptoms in line with a female gender role (often in tears, depressed mood), and that girls more often seek help from friends than boys do in adolescence. Research shows that, in general, the majority of boys also seek help. However, the abovementioned research is typically based on western cultures.

Another and related possibility may concern how the adolescent’s symptoms are categorized—by themselves—but also by parents, teachers, and clinicians. In Study I,

irritability and concentration difficulties were the most common symptoms of DSM-5-defined depression. It is possible that irritability is more often considered as part of conduct disorder (CD) in boys and depression in girls. In the same vein, concentration difficulties might more often be categorized as ADHD in boys and depression in girls (Mayes et al., 2019). In line with such a proposal is a study showing that girls more often receive treatment for emotional problems for much longer than boys before receiving a thorough assessment of neuropsychiatric symptoms (Klefsjö et al., 2021). Moreover, Fresson et al. (2019) found that after the male gender stereotype depicting boys as inattentive and impulsive was activated, the child's behavior was rated more negatively, and performance on a neuropsychiatric assessment declined if the rater thought that the child was a boy. These results may indicate that depression is more often neglected in boys. These issues need further investigation.

In conclusion, the gender difference in early adolescent depression might be somewhat inflated. However, evidence to date does not suggest that the female preponderance in depression is *completely* explained by such processes. Research should further investigate whether adolescents show evidence of response bias; whether boys are underdiagnosed with depression; to what extent these biases are driven by male and female gender roles; and how to avoid them.

4.5.3.2.The gender binary

Psychological research has traditionally operated with two genders: girls/women and boys/men. This includes the TESS study. Evidently, though, the issue of gender terminology is more complex, and the terminology is developing. Indeed, over the last two decades, increased recognition of the intersex (e.g., Dreger & Herndon, 2009) and transgender (e.g., Martínez-San Miguel & Tobias, 2016) rights movements have challenged the binary view of gender. Thus, since the initiation of TESS in 2007, the macrosystem—the values and sociocultural characteristics of the broader culture affecting the developing child's or

adolescent's identity, values, and perceptions (Bronfenbrenner, 1977)—has changed. The macrosystem influences the structures and activities occurring at the concrete level (Bronfenbrenner, 1977), e.g., measurement of gender in research.

A recent review shows how different empirical areas undermine the gender binary (Hyde et al., 2019). This includes findings from neuroscience. Even in brain-areas shown to have mean-level differences in structure and functioning, most people have a "mosaic" of more-or-less gender-typical organizations (i.e., there is no internal consistency in which women generally have a "female brain" and which men generally have a "male brain"). The mosaic metaphor also applies to the few psychological characteristics shown to have a mean gender difference. Even hormonal systems show considerable overlap (i.e., not gender dimorphic) (Hyde et al., 2019).

Relying exclusively on the gender binary in research excludes other important research questions. Studies have shown that people not identifying in the binary have less perceived social support and are at increased risk of bullying victimization (Aparicio-García et al., 2018); self-harm and suicide (Jadva et al., 2023); and anxiety and depression (Thorne et al., 2019). The unique challenges of those identifying as non-binary—including differentiating transgender/non-transgender and female/male assigned at birth (see e.g., Rimes et al., 2019)—needs continued attention.

Summing up, research on the female preponderance in depression is limited by the gender binary. This includes our Study III. Future studies may include children's and adolescents' self-reports of which gender they identify with (both binary and non-binary). This information may lead to a more nuanced view on the gender difference in depression, including if explanations for the female preponderance may extend to adolescents identifying as non-binary. Notably, current surveys in Norway indicate that 99.6% of adults consider themselves as either male or female—only 0.4 % as nonbinary (Statistics Norway, 2021).

Even though the estimates for adolescents might be higher, identifying as non-binary is likely rare. Thus, if a third non-binary category is to be included in quantitative research designs, the study needs to plan for a considerably larger sample size needs than the TESS-sample.

4.6.Methodological considerations

The studies in the current thesis are all based on data from the TESS, a large-scale longitudinal study. We applied multi-method and multi-informant assessment, limiting common-method effects that may otherwise have inflated the associations between study variables. Moreover, we used within-person methodology that accounts for time-invariant confounding effects. The results shed light on the stability of depression, the relationship between adolescent depression and personality, and possible explanations for the gender difference in adolescent depression. However, when interpreting these findings, several methodological issues should be considered.

4.6.1. The sample

4.6.1.1.Selection bias

An important question regarding the studies included in this thesis concerns their generalizability: whether results hold beyond the current study (Bordens & Abbott, 2022). *Selection bias* occurs when the selection of subjects into the study (sampling bias)—or their likelihood of remaining in the study (attrition)—leads to systematic differences in results because of a biased sample. The generalizability of the study thus decreases (e.g., Gerhard, 2008).

In the TESS, all children born in 2003 and 2004 were invited to participate. There was, however, likely some sampling bias from the non-participation at T1. 1) n = 176 families were excluded because of insufficiently proficiency in Norwegian. These may represent a more ethnically diverse subpopulation in Trondheim, possibly with lower SES. 2) n = 166 families were mistakenly not asked to participate by the health nurses, probably randomly

distributed, but we do not have data to confirm this. Finally, 3) n = 539 families declined to participate before the sample was drawn, and 4) 2.7% of the two cohorts that first were invited to participate (N = 3,456) did not show-up for the 4-year-old health check-up. We cannot exclude the possibility that parents in these families declined or did not show up because their child had mental health issues or because of other stressful circumstances. Notably, the TESS is largely representative of the Norwegian population regarding parents' level of education (Steinsbekk & Wichstrøm, 2018). It did, however, include more divorced parents (7.6%) compared to the general population (2.1%; Statistics Norway 2017). Moreover, other parts of Norway have more ethnic diversity. Thus, our findings may not be generalized to more diverse populations.

Attrition is a typical problem that influences the generalizability of longitudinal studies. Comparable studies on Norwegian population-based samples have shown attrition rates of 48% over 12 years (Sigurdson, 2019) and 56% over 15 years (Gustavson et al., 2012). As can be seen in Appendix Figure I: Flowchart of Sample Recruitment, the drop-out-rate in the TESS was closer to 30% over 12 years (T1 to T7). In all three studies, our attrition analyses showed that symptoms of MDD and dysthymia at certain measuring points predicted later drop-out, and in Study III, bullying victimization at age 6 predicted attrition at ages 10, 12 and 14. This suggests some systematic attrition and that we might have "lost" some of the depressed children and adolescents—especially from ages 12 to 14—as well as some of the 6-year-olds experiencing bullying victimization. However, these results could conceivably have been produced by the many attrition analyses conducted. To rule this out, in Studies II and III, we investigated whether the attrition was missing at random (MAR) or missing completely at random (MCAR). The results did suggest that the attrition was at least MAR, suggesting some systematic attrition. A limitation with Study I is that we only run ordinary regression analyses to check for systematic missing data and not the more formal Little's test.

Collectively, we had some sample bias and some attrition. Finally: Norway is a country with low rates of psychiatric disorders (Bøe et al., 2021; Wichstrøm et al., 2012), and may differ from contexts with overall lower and more variation in SES and more ethnic diversity, such as the United States. All these elements imply that the prevalence rates in our studies have somewhat limited generalizability. Even so, besides the prevalence of symptoms and disorders in Study I—and of stressors in Study III—the rest of our research questions were concerned with associations between variables. Associations first and foremost depend on the *variance* and not the mean levels of the included variables (Field, 2013). Moreover, a longitudinal Norwegian study investigating the influence of attrition on generalizability found evidence to suggest that even though attrition influences mean levels, estimates of associations are more robust (Gustavson et al., 2012). In conclusion, our results may be generalized to a reasonable degree—however with caution.

4.6.1.2.Sample size

When using within-person methodology, a large sample has been recommended (e.g., Masselink et al., 2018). With our sample sizes (Study I: n = 1,042, Study II: n = 817, Study III: n = 748), there were some research questions we were not positioned to investigate. In Study I (on stability), with a bigger sample, we could have investigated models separately for boys and girls. It is conceivable, for example, that there is an increased risk of recurrence for girls when compared to boys. Also, relevant for Study II—a gender difference in neuroticism has been reported to emerge during the early adolescent years: girls increasing while boys decreasing somewhat (Soto et al., 2011). Future research with a larger sample may examine whether the predisposition and scar-explanations for the Big Five personality traits differ among girls and boys. Furthermore, the five traits could have been examined in multivariate models. Even though we were able to account for the arguably most important trait, neuroticism (e.g., Goldstein et al., 2018), the research base on the Big Five and depression in

adolescence consists of few studies, and these have considerable methodological issues. Knowledge of which other traits might be important to account for, is therefore somewhat uncertain.

4.6.2. Confounding

In our studies, we investigated predictions from earlier depression on later depression (Study I); personality traits on later depression and the reverse (Study II) and from SLEs and bullying victimization on later depression (Study III). As noted previously, observational longitudinal studies, such as ours, are not positioned for causal inferences. This is partly because of omitted variables, or confounders; which are the common cause of two observed variables and can lead to a spurious correlation, or a suppression of their causal relationship (Hamaker et al., 2020).

Confounding effects may be time-invariant (i.e., factors with equal amount of influence on the predictor and depression at different time-points) or time-varying (i.e., factors with particularly strong impact on the predictor at certain time-points). In all studies, we used within-person methodology, accounting for time-invariant confounding effects potentially produced by unmeasured factors (Berry & Willoughby, 2017; Hamaker et al., 2020). However, time-*varying* confounding effects may still have confounded our results (e.g., Hamaker et al., 2020). Examples of these are pubertal hormones first increasing and exerting their effects on depression in early adolescence (Stumper & Alloy, 2021) or peer stress, shown to predict depression more strongly with increasing age during adolescence (Hankin et al., 2015). In Study II, we included gender and neuroticism as observed variables that may have time-varying confounding effects. However, in all three studies, other unmeasured variables may also have had time-varying confounding effects on the results.

4.6.3. Measurement

4.6.3.1.Depression

In the current work, we used standardized semi-structured clinical interviews to measure symptoms of MDD and dysthymia: the Preschool Age Psychiatric Assessment (PAPA), the Child and Adolescent Psychiatric Assessment (CAPA), and the Schedule for Affective Disorders and Schizophrenia for School-Age Children (Kiddie-SADS). There are important advantages to using such interviews (Essau & Ollendick, 2009). First, when compared to both questionnaires—which usually do not cover all the DSM-defined symptoms—and unstructured interviews—semi-structured interviews are designed to accurately assess diagnostic criteria and are therefore often better equipped to assess diagnoses (Mueller & Segal, 2014). Also, when compared to questionnaires specifically: interviews increase the likelihood of capturing symptoms that may be more difficult for the child or adolescent to notice if they are depressed. These are all elements that improve the validity of the depression measure—the extent to which an instrument is measuring what it is intended to measure (Carmines & Zeller, 1979). Further on, when compared to unstructured interviews, semi-structured interviews reduce variability among those interviewing (Mueller & Segal, 2014), and there is less risk for different interviewers to ascribe different degrees of clinical significance to particular symptoms (Le Couteur & Gardner, 2008). Both of these aspects increase inter-rater reliability, which is a prerequisite for validity (Mueller & Segal, 2014). Finally, semi-structured interviews also have an important advantage when compared to *structured* interviews in that they include the possibility for the interviewer to follow up with probes to more accurately assess whether a symptom is present or not. As such, they increase the likelihood that interviewers do not prematurely conclude that a diagnostic criterion is not fulfilled and move on (Mueller & Segal, 2014). However, semi-structured interviews therefore require more clinical judgment than structured interviews (Mueller &

Segal, 2014). In the TESS, the interviewers had at least a bachelor's degree in a relevant field, relevant experience, and received training from the team that developed PAPA/CAPA, and with a psychologists or psychiatrist when learning how to administer and score the Kiddie-SADS. The interviewers' experiences with the use of Kiddie-SADS were also continually discussed in the research group's meetings (every second week).

According to Achenbach (2006), multi-informant data is required for a comprehensive assessment of psychopathology in children, because different informants contribute with unique information about an individual's functioning. Following this logic, we interviewed both the children/adolescents and their parents separately. A symptom was considered present if it was endorsed by either informant. We were thus able to measure symptoms that likely require the introspection of the child/adolescent, for example, low self-esteem or suicidal thoughts. Other symptoms may be more available to the parents, such as when a child suddenly becomes more socially withdrawn with siblings or at the dinner table—a change that the child might not notice themselves. Thus, interviewing multiple informants increases the overall validity—in our case—the likelihood of capturing all the DSM-5-defined depressive symptoms of MDD and dysthymia.

Despite the advantages of multi-informant, semi-structured clinical interviews to measure depression in children and adolescents, there are several limitations. In a review, Duncan et al. (2019) found vast variations in the reliability of clinical interviews—and thus also lower validity (De Los Reyes et al., 2015; Mueller & Segal, 2014). Of note, in our sample, the PAPA, the CAPA, and the Kiddie-SADS showed evidence of adequate ICC inter-rater reliability (Koo & Li, 2016). More directly related to validity: research has shown that misunderstandings may occur during interviews (e.g., Kessler & Üstün, 2004). Moreover, there are challenges in measuring psychopathology in children and adolescents through the use of interviews, even though these arguably also apply to questionnaires. Weiss and Garber

(2003) propose that young children may lack considerable abilities for introspection and conceptualization of symptoms, and therefore also the ability to report on these symptoms—including to their parents. In that sense, parents may be better informants of *adolescents'* rather than children's depressive symptoms, at least those that require introspection (e.g., low self-esteem). However, adolescents' normative social reorientation toward peers may entail that they share *less* with their parents. Thus, how informative self- versus parent-reports are may vary depending on the developmental stage of the child or adolescents. These developmental differences underscore the advantages of using multi-informant data to increase the overall likelihood of capturing symptoms.

As previously noted, at age 16, we changed the measurement of depression from CAPA to Kiddie-SADS. For a symptom of depression to be coded as present with the Kiddie-SADS, the requirements are stricter in that the symptom must be present for most of the day/at least 50% of the day—in line with the formulation in the DSM-5. However, in the CAPA, most of the symptoms are scored affirmatively if present for at least one hour of the day. One possibility is that the CAPA duration-criterion is too lenient. However, in Study I, we found evidence indicating that we measured the same MDD and dysthymia constructs across ages 4 to 14, suggesting that CAPA *did* capture DSM-5-defined depression in children and adolescents. Alternatively, the Kiddie-SADS "most of the day" criterion—which aligns with the DSM-5-duration criterion—might be too strict and lead to more false negatives (i.e., not registering "real" depressive symptoms) in our sample at age 16. It might be that a one-hour criterion is better suited for children and adolescents. To the best of my knowledge, previous studies have not investigated whether using the CAPA compared to the Kiddie-SADS provides different DSM-5-defined depression prevalence rates—and if so, why. The CAPA has, however been compared to the clinical interview Development and Well-Being Assessment (DAWBA), and results revealed that by applying the CAPA, twice as many

diagnoses were obtained as when the DAWBA was used (Angold et al., 2012), indicative of the CAPA not being a strict diagnostic instrument.

In Studies II and III, depressive symptoms were captured by a collective measure of symptoms of MDD and dysthymia. These conditions overlap substantially (see Table 1). However, there is evidence suggesting that predictors for MDD and dysthymia might differ (Markkula et al., 2017). Thus, it is possible that the results of the studies might have diverged somewhat if the conditions had been investigated separately. For example, Schramm et al. (2020) have reviewed research on adults, showing that dysthymia is more strongly correlated with higher levels of neuroticism and lower levels of extraversion than non-chronic MDD.

From March 2020, the COVID-19 pandemic led to various lockdowns and social restrictions. The T7 data collection had just begun (28th of January 2020) and ran until January 7, 2022. Thus, most of our latest data collection (included in Study II) was performed during the pandemic. When restrictions were at their strictest, adolescents were not able to freely socialize, explore, and interact in person with their friends—in this developmental period marked by a social orientation towards peers. Social isolation and loneliness are important predictors of depression (Dunn & Sicouri, 2022). Moreover, considerable research has shown that the pandemic might have had negative consequences for adolescents' mental health, including increases in depressive symptoms (Ludwig-Walz et al., 2022). Four things to note regarding our sample at T7. First, the data-collection spread out in time such that a few adolescents were interviewed just prior to the pandemic, and some also by the end (i.e., fewer restrictions). Second, during the data-collection, the restrictions varied greatly, with the strictest restrictions lasting only a few weeks at a time. Third: variations in restrictions in schools and life circumstances (e.g., if the adolescent or someone they lived with had an underlying somatic condition) mean that the pandemic likely affected the adolescents differently. These differences in restrictions over time—and between participants—might

have influenced our data and thus the results of Study II. Importantly, there were far fewer restrictions in Trondheim than in other parts of the country, such as Oslo.

A final point regarding how we measured and analyzed depression regards the issue of comorbidity. Depression in childhood and adolescence has shown high comorbidity—especially with anxiety disorders (Sharma et al., 2019; Vasileva et al., 2021) but also with ADHD (Sandstrom et al., 2021), and disruptive mood dysregulation disorder (DMDD)—at least in preschool, when DMDD is most common (Copeland et al., 2013b). It is possible that results in our studies would have differed for children with comorbidities when compared to those with "pure" depression diagnoses. This may be considered in future investigations.

4.6.3.2.Personality

In Study II, personality traits were measured across ages 10 to 16 with the Big Five Inventory (BFI-46A; Soto et al., 2008). This questionnaire captures the five traits described in the Five Factor Model (FMM): neuroticism, extraversion, conscientiousness, agreeableness, and openness. The reliability was lower at age 10 (α=.59-.69 for the different traits) than at ages 12, 14 and 16 (α=.67-.83)—comparable to a previous study also showing higher internal consistency in older age-groups (Soto et al., 2008). The lower reliability in the 10-year-olds increases the risk of both Type 2-errors in the youngest age-span. Even so, Soto et al. (2008) also found that when accounting for acquiescence (some tend to agree a lot, some tend to agree less), the Big Five structure was clearly recognizable in the 10-year-olds, and with congruence coefficients of at least 0.93 for each of the five components.

In our study, we measured the five higher-order traits of the FMM. However, these traits include different facets. The exact number and names of the facets depend on the questionnaire. In the BFI-46A, 10 facets have been identified: Anxiety and Depression (neuroticism), Activity and Assertiveness (extraversion), Order and Self-Discipline (conscientiousness), Compliance and Altruism (agreeableness) and Aesthetics and Ideas

(openness) (Soto & John, 2009). First, it is possible that some of the facets drive the predictions of the overall traits of depression. For example, Goldstein et al. (2018) found that only the depression facet—and not the anxiety facet—of neuroticism predicted the first onset of depression in a sample of adolescent girls. A second issue is whether combinations of different facets—and/or combinations of the five overall traits—might confer an increased risk for depression, over and above the facets/traits alone. For example, Allen et al. (2018) found a three-way interaction in adults in which those scoring high on neuroticism, low on extraversion, and low on conscientiousness were those most likely to score high on depression. Moreover, they found that this interaction effect was driven by the facets of withdrawal (neuroticism), enthusiasm (extraversion) and industriousness (conscientiousness). However, a study on temperamental traits found solid evidence for a three-way interaction in that higher negative emotionality, coupled with lower positive emotionality and lower effortful control, was most strongly associated with depressive symptoms (Vasey et al., 2013). To my knowledge, whether such three-way interactions at the trait and facet levels of the Big Five personality traits may also be identified in adolescents has not been investigated.

4.6.3.3.Stressful life events

In Study III, stressful life events were assessed by a self- and parent-reported checklist across ages 8 to 14 (see Appendix). The checklist covers important life events (e.g., new sibling, parental divorce) and very serious life events (been in an accident and almost died, sexual assault). Several issues with this measure deserve attention. First, as with all self-report questionnaires, the informants' developmental level might influence the validity of the measure. The checklist was completed while the research assistant was present, and in the youngest age-groups (age 8 and 10) it was made explicit to the child/adolescent to ask questions if they did not understand the formulation. Even so, there is never a guarantee that they asked. Secondly, the broad range of seriousness in the checklist makes it unclear whether

the results shown in Study III stem from the most serious life events. However, we did explore this possibility and found that predictions from important versus more serious events on depression did not differ. Finally, the events may vary as to how readily the child or adolescent might endorse the item. Some events, such as sexual abuse, likely elicit more shame (e.g., Pettersen, 2013) than quarreling with a friend and are thus less likely to be reported. If the information on some events is more reliable than on others, this limits the validity of the measure.

In the research on SLEs, some use interviews (e.g., Hammen, 2005). These have several advantages, such as the possibility of discerning the *meaning* of an event for the adolescent, including whether the event was experienced as stressful. For example, even though parental divorce is associated with adolescent depression at the group level (Liu, 2022), for some adolescents, the divorce may provide relief and closure after years of conflicts between the parents. Some interviews also allow one to discern whether a stressor was dependent or independent on the informants' behavior (Hammen, 2005), and whether it was perceived as interpersonal or not. This information could give a more nuanced understanding of the role of stress in the gender difference in depression (Hankin et al., 2007; Shih et al., 2006). Ideally, a study may combine questionnaires and interviews (see for example Hankin et al., 2007).

4.6.3.4.Bullying victimization

Finally, in Study III, bullying victimization was measured with the Olweus Bully Victim Questionnaire (OBVQ; Solberg & Olweus, 2003) teacher report version. The measure is based on the following definition of bullying victimization: the intention to harm the victim, the repetitive nature of bullying, and finally: a power imbalance between the bully(ies) and the victim (Solberg & Olweus, 2003). It includes items on relational and physical bullying, and in our study, we used this questionnaire to capture bullying victimization

overall. There are several advantages to using teacher reports. First of all, for many adolescents, being a victim of bullying may be associated with shame (Strøm et al., 2018). This might decrease the likelihood of them admitting to being victims—even in an anonymous research project. In that sense, using teacher reports might increase the validity of the measure. Secondly, given the amount of time children and adolescents spend in school, the likelihood of teachers becoming aware of bullying victimization might be greater than if we were to use parent-reports.

There are, however, several limitations to our measurement. For example, the power imbalance entailed in the bullying victimization dynamic may not only play in front of the teacher—but also in breaks without supervision, and on social media. Moreover, some types of relational bullying—such as rolling of the eyes in a specific context or spreading of rumors—might be harder for a teacher to detect than physical bullying victimization. To gain more knowledge, peer-reports may also be used. These have, however, been shown to provide substantially lower rates of bullying victimization than self-reports (Baly et al., 2014; Branson & Cornell, 2009). Self-reports, on the other hand, may be subject to inflation, in those cases where the adolescent does *not* incorporate the important concept of power-imbalance when they report on bullying victimization events (Baly et al., 2014). Future research may consider using a combined measure of self-, peer-, and teacher-reports.

Further on, we investigated bullying overall. However, different types of bullying victimization (relational versus physical) may have different effects. For example, one meta-analysis found a stronger association between relational and internalizing problems than between physical victimization and internalizing problems (Casper & Card, 2017). Of note, the same meta-analysis did find that both associations were of moderate strength and that the inter-correlations (suggesting overlap in constructs) between physical and relational bullying victimization were strong (Casper & Card, 2017). Moreover, others have found that physical,

relational, and verbal bullying victimization predict later depression in adolescents to an equal degree (Klomek et al., 2019). Arguably, different types of bullying victimization hurt the fundamental need to belong (Baumeister & Leary, 1995). Unmet belongingness needs have been shown to increase the risk for depression (Verhagen et al., 2018). Even so—future studies may aim to include more thorough measures of both types of bullying victimization and investigate potential differing roles in the female preponderance in depression.

In Study III, we did not have a measure for cybervictimization, an important risk-factor for adolescent depression (Hu et al., 2021). In fact, cyberbullying victimization has been suggested to have even more long-lasting and detrimental effects than traditional bullying because bullying material may be stored online (Van Geel et al., 2014). The TESS was planned and first executed in 2007. The TESS not including a measure for cybervictimization must be seen in the context of the fact that the use of smart phones and social media first dramatically increased around 2011-2012 (Statista, 2023). Future research on specific types of bullying should consider including a measure on cyberbullying.

4.7.Implications and suggestions for future research

The current thesis focuses on the stability of depression from preschool into adolescence, the relationship between depression and personality in adolescence, and the female preponderance in adolescent depression. The studies included were based on multi-informant clinical interviews of symptoms of MDD and dysthymia, and predictions were examined at a within-person level. Prior research on the specific research questions of this thesis has mainly used methodologies that conflates between- and within-person information (e.g., CLPMs). Moreover, the baseline ages of these studies are typically 12-13 years, paralleling the early adolescent increase—and the emerging female preponderance—in depression. Thus, to further illuminate which risk and vulnerability factors may contribute to explaining early adolescent depression, we need studies with younger baseline ages.

Our findings solidify prior research indicating the earliest adolescent years as a particular period of risk for depression. First, all five types of stability that we investigated increased (Study I). Second, even though the studies included in the current thesis are not positioned to uncover causal relations, we found evidence to suggest that depressive symptoms may impact personality traits already from early adolescence on: increased neuroticism and decreased extraversion and conscientiousness (Study II). Finally, we replicated the female preponderance in early adolescent depression (Study III). Collectively, these results underscore the importance of preventative efforts in early adolescence—aimed at reducing the increase in depression and its potential negative influence on personality development.

Knowledge of risk factors is important in order to develop effective preventative interventions. Our findings are in line with proposals that earlier depression (Study I) and increased neuroticism (Study II) confer vulnerability for depression in early adolescence. Furthermore, early adolescent girls may react to SLEs, and bullying victimization with more depressive symptoms than early adolescent boys (Study III). Thus, depression in earlier childhood, neuroticism, and SLEs and bullying victimization may be considered in preventative efforts. Importantly, our findings suggest that these risk factors are already at play already at age 10. Thus, it is likely beneficial to instigate preventative efforts at least by this age, and not wait until, e.g., the adolescents are in high school and 12 to 13 years old. Regarding which concrete mechanisms interventionalists should target, prior research alludes to some possibilities. For example, depressive symptoms might increase the likelihood for sadness rumination (Spyropoulou & Giovazolias, 2022), excessive reassurance seeking (Starr & Davila, 2008), and co-rumination (e.g., Rose, 2021). These may all potentially lead to social erosion or other forms of interpersonal stress—which thereafter may increase the risk of recurrence or worsening of the depressive symptoms (e.g., Hankin et al., 2010; Rose et al.,

2017). Adolescents scoring particularly high on neuroticism may also be especially vulnerable because they respond to stress with more self-blame (Liu et al., 2020) and emotional suppression (Yoon et al., 2013), or even *perceive* stress more readily (Hoferitcher et al., 2014). Thus, in both preventative efforts targeting subthreshold depressive symptoms, as well as in treatment, interventionalists could focus on targeting potential tendencies to excessive reassurance seeking, rumination, self-blame, and emotional suppression. Attention should also be paid to reactions after an adolescent has experienced SLEs or bullying victimization. Girls may also have an increased "depressogenic vulnerability" (Hyde et al., 2020), involving, for example, higher levels of body dissatisfaction (von Soest, et al. 2016) and self-criticism in general (Xavier et al., 2016) than boys, and thereby develop depressive symptoms when faced with stress. Social comparison on social media for may be important to target in this regard. It is also possible that after exposure to these stressors, girls engage in more rumination (Hampel & Peterman, 2005), and co-rumination (Rose et al., 2017; Stone et al., 2011). Notably, though, most of the studies indicating potential mechanisms are cross-sectional, have baseline ages in middle adolescence, and/or are conducted on adults. More research should be done on the early adolescent period specifically, preferably already from age 10 and onwards.

In both preventive efforts and the assessment of need for treatment, it is important to detect those children or adolescents that have already developed depressive symptoms. In this regard, Study I has three potential implications. First, our findings indicate that the DSM-5 MDD and dysthymia may be used in diagnosing DSM-defined depression in children and adolescents. Notably, we did not address the developmental *appropriateness* of the DSM-5, i.e., whether other and/or additional symptoms represent these depression constructs in children and adolescents. Second, the findings suggest that the likelihood that a single depressive symptom is accompanied by several others increases in adolescence. Third, the most common symptoms were irritability and concentration difficulties, while some of the

least common symptoms were persistent depressed mood and/or anhedonia—the core symptoms (must be present) for MDD and dysthymia in adults. The ICD-11 does not include the DSM-5 age-adjustment that includes irritability as a core symptom in children and adolescents. Therefore, if clinicians in Norway—and other countries using the ICD-11 diagnostic system—strictly follow the ICD-11 criteria, they will miss depressive episodes in which irritability and not the other core symptoms are present.

4.8.Conclusions

The overarching research questions (Page 5) can be answered as follows:

a) From preschool age into adolescence, within-person increases or decreases in symptoms in DSM-defined major depressive disorder (MDD) and dysthymia predicted corresponding within-person changes. These findings are in line with the proposal that the risk of recurrence occurs because depression changes characteristics in the child/adolescent or their environments, which thereafter increases the risk of depression. Preventative and treatment interventions may consider targeting potential "depressive scars". Further on, we found modest to moderate rank-order stability (relative to others depression levels), higher absolute stability (relative to ones' own depression level), and evidence that the same DSM-defined MDD and dysthymia constructs were measured across preschool until adolescence. Thus, our results indicate the importance of preventing depression from an early, particularly in early adolescence. The most typical symptoms were irritability and concentrating difficulties, which may be relevant when identifying depression across preschool until adolescence. All types of stability increased when entering adolescence, when the prevalence also increased.

b) Increases in neuroticism predicted increases in depressive symptoms, in line with the proposal that neuroticism may be a vulnerability factor for depression. This finding may be accounted for in preventative interventions. Moreover, adolescent depression might

have an even wider impact on personality development: increased depressive symptoms predicted increased neuroticism and decreased extraversion and conscientiousness—underscoring the importance of preventing adolescent depression.

c) When entering adolescence, we found evidence suggesting that girls react more strongly to stressful life events and bullying victimization than boys in the form of more depression—in line with a stress reactivity explanation for the female preponderance in depression. A potential increased stress reactivity in girls may be accounted for in preventative efforts targeting adolescents exposed to stress.

References

Achenbach, T. M. (2006). As others see us: Clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science, 15*(2), 94-98. https://doi.org/10.1111/j.0963-7214.2006.00414

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213-232. https://doi.org/10.1037/0033-2909.101.2.213

Adams, M. J., Howard, D. M., Luciano, M., Clarke, T.-K., Davies, G., Hill, W. D., 23andMe Research Team, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Smith, D., Deary, I. J., Porteous, D. J., & McIntosh, A. M. (2020). Genetic stratification of depression by neuroticism: revisiting a diagnostic tradition. *Psychological Medicine, 50*(15), 2526-2535. https://doi.org/10.1017/S0033291719002629

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review, 30*(2), 217-237. https://doi.org/10.1016/j.cpr.2009.11.004

Allen, T. A., Carey, B. E., McBride, C., Bagby, R. M., DeYoung, C. G., & Quilty, L. C. (2018). Big Five aspects of personality interact to predict depression. *Journal of Personality, 86*(4), 714-725. https://doi.org/10.1111/jopy.12352

Allport, G. W. (1961). *Pattern and growth in personality.* Holt, Rinehart and Winston.

American Psychiatric Association, A. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Publications.

American Psychological Association, A. (2020). *Publication manual of the American psychological association (7th ed.)*. American Psychological Association.

Angold, A., & Costello, E. J. (2000). The child and adolescent psychiatric assessment (CAPA). *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(1), 39-48. https://doi.org/10.1097/00004583-200001000-00015

Angold, A., Erkanli, A., Copeland, W., Goodman, R., Fisher, P. W., & Costello, E. J. (2012). Psychiatric diagnostic interviews for children and adolescents: a comparative study. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(5), 506-517. https://doi.org/10.1016/j.jaac.2012.02.020

Aparicio-García, M. E., Díaz-Ramiro, E. M., Rubio-Valdehita, S., López-Núñez, M. I., & García-Nieto, I. (2018). Health and well-being of cisgender, transgender and non-binary young people. *International Journal of Environmental Research and Public Health, 15*(10), 1-11. https://doi.org/10.3390/ijerph15102133

Arain, M., Haque, M., Johal, L., Mathur, P., Nel, W., Rais, A., Sandhu, R., & Sharma, S. (2013). Maturation of the adolescent brain. *Neuropsychiatric Disease and Treatment*, 449-461. https://doi.org/10.2147/NDT.S39776

Avenevoli, S., Swendsen, J., He, J.-P., Burstein, M., & Merikangas, K. R. (2015). Major depression in the national comorbidity survey–adolescent supplement: Prevalence, correlates, and treatment. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(1), 37-44. https://doi.org/10.1016/j.jaac.2014.10.010

Bailen, N. H., Green, L. M., & Thompson, R. J. (2019). Understanding emotion in adolescents: A review of emotional frequency, intensity, instability, and clarity. *Emotion Review, 11*(1), 63-73. https://doi.org/10.1177/17540739187688

Bakken, A. (2020). *Ungdata 2020. Nasjonale resultater*. NOVA Rapport 16/20. NOVA, OsloMet. https://www.forebygging.no/globalassets/ungdata-2020-nasjonale-resultater-nova-rapport-16-20-1.pdf

Baly, M. W., Cornell, D. G., & Lovegrove, P. (2014). A longitudinal investigation of self-and peer reports of bullying victimization across middle school. *Psychology in the Schools, 51*(3), 217-240. https://doi.org/10.1002/pits.21747

Barańczuk, U. (2019). The Five Factor Model of personality and social support: A meta-analysis. *Journal of Research in Personality, 81*, 38-46. https://doi.org/10.1016/j.jrp.2019.05.002

Barzeva, S. A., Richards, J. S., Veenstra, R., Meeus, W. H., & Oldehinkel, A. J. (2022). Quality over quantity: A transactional model of social withdrawal and friendship development in late adolescence. *Social Development, 31*(1), 126-146. https://doi.org/10.1111/sode.12530

Bastin, M., Mezulis, A. H., Ahles, J., Raes, F., & Bijttebier, P. (2015). Moderating effects of brooding and co-rumination on the relationship between stress and depressive symptoms in early adolescence: A multi-wave study. *Journal of Abnormal Child Psychology, 43*, 607-618. https://doi.org/10.1007/s10802-014-9912-7

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497-529. https://doi.org/10.1037/0033-2909.117.3.497

Bråten, R.H., Sten-Gahmberg, S., Bugge, C., Kristiansen, I.S., Gardsjord, E.S., Sæther, E.M., 2023. Økt bruk av psykiske diagnoser for unge. *Journal of the Norwegian Psychology Association, 60*(4)*,* 208–215. https://doi.org/10.52734/vhhp4493

Bernaras, E., Jaureguizar, J., & Garaigordobil, M. (2019). Child and adolescent depression: A review of theories, evaluation instruments, prevention programs, and treatments. *Frontiers in Psychology, 10*(543), 1-24. https://doi.org/10.3389/fpsyg.2019.00543

Berry, D., & Willoughby, M. T. (2017). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child Development, 88*(4), 1186-1206. https://doi.org/10.1111/cdev.12660

Bertha, E. A., & Balázs, J. (2013). Subthreshold depression in adolescence: a systematic review. *European Child & Adolescent Psychiatry, 22*(10), 589-603. https://doi.org/10.1007/s00787-013-0411-0

Blomqvist, K., Saarento-Zaprudin, S., & Salmivalli, C. (2020). Telling adults about one's plight as a victim of bullying: Student-and context-related factors predicting disclosure. *Scandinavian Journal of Psychology, 61*(1), 151-159. https://doi.org/10.1111/sjop.12521

Boeninger, D. K., Masyn, K. E., Feldman, B. J., & Conger, R. D. (2010). Sex differences in developmental trends of suicide ideation, plans, and attempts among European American adolescents. *Suicide and Life-Threatening Behavior, 40*(5), 451-464. https://doi.org/10.1521/suli.2010.40.5.451

Bond, L., Carlin, J. B., Thomas, L., Rubin, K., & Patton, G. (2001). Does bullying cause emotional problems? A prospective study of young teenagers. *Bmj, 323*(7311), 480-484. https://doi.org/10.1136/bmj.323.7311.480

Bordens, K., & Abbott, B. B. (2022). *Research Design and Methods: A Process Approach* (10th ed.). McGraw-Hill Education.

Bordini, B., & Rosenfield, R. L. (2011). Normal pubertal development: Part I: The Endocrine Basis of Puberty. *Pediatrics in Review, 32*(6), 223-229. https://doi.org/10.1542/pir.32-6-223

Bouma, E. M., Riese, H., Ormel, J., Verhulst, F. C., & Oldehinkel, A. J. (2009). Adolescents' cortisol responses to awakening and social stress; effects of gender, menstrual phase and oral contraceptives. The TRAILS study. *Psychoneuroendocrinology, 34*(6), 884-893. https://doi.org/10.1016/j.psyneuen.2009.01.003

Bowlby, J. (1982). *Attachment and loss: Attachment* (2nd ed, vol. 1). Basic books.

Branje, S. J., Hale, W. W., Frijns, T., & Meeus, W. H. (2010). Longitudinal associations between perceived parent-child relationship quality and depressive symptoms in adolescence. *Journal of Abnormal Child Psychology, 38*, 751-763. https://doi.org/10.1007/s10802-010-9401-6

Branson, C. E., & Cornell, D. G. (2009). A comparison of self and peer reports in the assessment of middle school bullying. *Journal of Applied School Psychology, 25*(1), 5-27. https://doi.org/10.1080/15377900802484133

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*(7), 513-531. https://doi.org/10.1037/0003-066X.32.7.513

Bruni, O., Sette, S., Fontanesi, L., Baiocco, R., Laghi, F., & Baumgartner, E. (2015). Technology use and sleep quality in preadolescence and adolescence. *Journal of Clinical Sleep Medicine, 11*(12), 1433-1441. https://doi.org/10.5664/jcsm.5282

Bruserud, I. S., Roelants, M., Oehme, N. H. B., Madsen, A., Eide, G. E., Bjerknes, R., Rosendahl. K., & Juliusson, P. B. (2020). References for ultrasound staging of breast maturation, Tanner breast staging, pubic hair, and menarche in Norwegian girls. *The Journal of Clinical Endocrinology & Metabolism, 105*(5), 1599-1607. https://doi.org/10.1210/clinem/dgaa107

, C. (2020). Family processes and children's and adolescents' well-being. *Journal of Marriage and Family*, *82*(1), 145-174. https://doi.org/10.1111/jomf.12637

Burcusa, S. L., & Iacono, W. G. (2007). Risk for recurrence in depression. *Clinical Psychology Review, 27*(8), 959-985. https://doi.org/10.1016/j.cpr.2007.02.005

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*(2), 261-304. https://doi.org/10.1177/0049124104268644

Büttner, C. M., & Rudert, S. C. (2022). Why didn't you tag me?!: Social exclusion from Instagram posts hurts, especially those with a high need to belong. *Computers in Human Behavior, 127*, 1-16. https://doi.org/10.1016/j.chb.2021.107062

Bøe, T., Heiervang, E. R., Stormark, K. M., Lundervold, A. J., & Hysing, M. (2021). Prevalence of psychiatric disorders in Norwegian 10-14-year-olds: Results from a cross-sectional study. *PLOS ONE, 16*(3), 1-15. https://doi.org/10.1371/journal.pone.0248864

Calvete, E., Orue, I., & Gamez-Guadix, M. (2016). Do extraversion and neuroticism moderate the association between bullying victimization and internalizing symptoms? A three-

wave longitudinal study. *Journal of School Psychology, 56*, 1-11. https://doi.org/10.1016/j.jsp.2016.02.003

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage Publications.

Carr, A. (2015). *The handbook of child and adolescent clinical psychology: A contextual approach*. Routledge.

Carskadon, M. A. (2011). Sleep in adolescents: the perfect storm. *Pediatric Clinics, 58*(3), 637-647. https://doi.org/10.1016/j.pcl.2011.03.003

Casper, D. M., & Card, N. A. (2017). Overt and relational victimization: A meta-analytic review of their overlap and associations with social–psychological adjustment. *Child Development, 88*(2), 466-483. https://doi.org/10.1111/cdev.12621

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychoogy, 56*, 453-484. https://doi.org/10.1146/annurev.psych.55.090902.141913

Caspi, A., & Shiner, R. (2006). Personality development. In N. Eisenberg (Ed.). *Handbook of child psychology* (6th ed., pp. 300-365). John Wiley & Sons, Inc.

Caspi, A., & Shiner, R. (2008). Temperament and personality. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor & A. Thapar (Eds.), *Rutter's child and adolescent psychiatry* (pp. 182–198). Wiley Blackwell.

Cavanagh, A., Wilson, C. J., Kavanagh, D. J., & Caputi, P. (2017). Differences in the expression of symptoms in men versus women with depression: a systematic review and meta-analysis. *Harvard Review of Psychiatry, 25*(1), 29-38.

Chan, S. W., Harmer, C. J., Norbury, R., O'Sullivan, U., Goodwin, G. M., & Portella, M. J. (2016). Hippocampal volume in vulnerability and resilience to depression. *Journal of Affective Disorders, 189*, 199-202. https://doi.org/10.1097/HRP.0000000000000128

Chang, C.-W., Yuan, R., & Chen, J.-K. (2018). Social support and depression among Chinese adolescents: The mediating roles of self-esteem and self-efficacy. *Children and Youth Services Review, 88*, 128-134. https://doi.org/10.1016/j.childyouth.2018.09.001

Chen, M. C., Hamilton, J. P., & Gotlib, I. H. (2010). Decreased hippocampal volume in healthy girls at risk of depression. *Archives of General Psychiatry, 67*(3), 270-276. https://doi.org/10.1001/archgenpsychiatry.2009.202

Chouinard, M. M., Harris, P. L., & Maratsos, M. P. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, 1-129. https://doi.org/10.1111/j.1540-5834.2007.00412.x.

Christina, S., Magson, N. R., Kakar, V., & Rapee, R. M. (2021). The bidirectional relationships between peer victimization and internalizing problems in school-aged children: an updated systematic review and meta-analysis. *Clinical Psychology Review*, *85*(101979), 1-19. https://doi.org/10.1016/j.cpr.2021.101979

Cicchetti, D., & Toth, S. L. (2009). A developmental psychopathology perspective on adolescent depression. In S. Nolen-Hoeksema & L. M. Hilt (Eds.), *Handbook of depression in adolescents* (pp. 1-31). Routledge/Taylor & Francis Group.

Clara, I. P., Cox, B. J., & Enns, M. W. (2003). Hierarchical models of personality and psychopathology: The case of self-criticism, neuroticism, and depression. *Personality and Individual Differences, 35*(1), 91-99. https://doi.org/10.1016/S0191-8869(02)00143-5

Clarke, T.-K., Zeng, Y., Navrady, L., Xia, C., Haley, C., Campbell, A., Navarro P., Amador, C., Adams, M. J., Howards, D. M., Soler, A., Hayward, C., Thomson, P. A., Smith, B. H., Padmanabhan, S., Hocking, L. J., Hall, L. S., Porteous, D. J., Major Depressive Disorder Working Group of the Psychiatric Genomic Consortium, . . . Howard, D. M. (2018). Genetic and environmental determinants of stressful life events and their overlap with depression and neuroticism. *Wellcome Open Research, 3*(11), 1-17. https://doi.org/10.12688/wellcomeopenres.13893.2

Clayborne, Z. M., Varin, M., & Colman, I. (2019). Systematic review and meta-analysis: adolescent depression and long-term psychosocial outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 58*(1), 72-79. https://doi.org/10.1016/j.jaac.2018.07.896

Compas, B. E., Jaser, S. S., Bettis, A. H., Watson, K. H., Gruhn, M. A., Dunbar, J. P., Williams, E., & Thigpen, J. C. (2017). Coping, emotion regulation, and psychopathology in childhood and adolescence: A meta-analysis and narrative review. *Psychological Bulletin, 143*(9), 939-991. https://doi.org/10.1037/bul0000110

Compian, L. J., Gowen, L. K., & Hayward, C. (2009). The interactive effects of puberty and peer victimization on weight concerns and depression symptoms among early adolescent girls. *The Journal of Early Adolescence, 29*(3), 357-375. https://doi.org/10.1177/027243160832365

Copeland, W. E., Adair, C. E., Smetanin, P., Stiff, D., Briante, C., Colman, I., Fergusson, D., Horwood, J., Poulton, R., Costello, Costello, J. E., & Angold, A. (2013a). Diagnostic transitions from childhood to adolescence to early adulthood. *Journal of Child Psychology and Psychiatry, 54*(7), 791-799. https://doi.org/10.1111/jcpp.12062

Copeland, W. E., Alaie, I., Jonsson, U., & Shanahan, L. (2021). Associations of childhood and adolescent depression with adult psychiatric and functional outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 60*(5), 604-611. https://doi.org/10.1016/j.jaac.2020.07.895

Copeland, W. E., Angold, A., Costello, E. J., & Egger, H. (2013b). Prevalence, comorbidity, and correlates of DSM-5 proposed disruptive mood dysregulation disorder. *American Journal of Psychiatry, 170*(2), 173-179. https://doi.org/10.1176/appi.ajp.2012.12010132

Copeland, W. E., Worthman, C., Shanahan, L., Costello, E. J., & Angold, A. (2019). Early pubertal timing and testosterone associated with higher levels of adolescent depression in girls. *Journal of the American Academy of Child & Adolescent Psychiatry, 58*(12), 1197-1206. https://doi.org/10.1016/j.jaac.2019.02.007

Cortes, K. I., & Kochenderfer-Ladd, B. (2014). To tell or not to tell: What influences children's decisions to report bullying to their teachers? *School Psychology Quarterly, 29*(3), 336-348. https://doi.org/10.1037/spq0000078

Costello, E. J., & Angold, A. (1988). Scales to assess child and adolescent depression: checklists, screens, and nets. *Journal of the American Academy of Child & Adolescent Psychiatry, 27*(6), 726-737. https://doi.org/10.1097/00004583-198811000-00011

Costello, E. J., Angold, A., Burns, B. J., Stangl, D. K., Tweed, D. L., Erkanli, A., & Worthman, C. M. (1996). The Great Smoky Mountains Study of Youth: goals, design, methods, and the prevalence of DSM-III-R disorders. *Archives of General Psychiatry, 53*(12), 1129-1136. https://doi.org/10.1001/archpsyc.1996.01830120067012

Cunningham, S., Hudson, C. C., & Harkness, K. (2021). Social media and depression symptoms: a meta-analysis. *Research on Child and Adolescent Psychopathology, 49*, 241-253. https://doi.org/10.1007/s10802-020-00715-7

Cyranowski, J. M., Frank, E., Young, E., & Shear, M. K. (2000). Adolescent onset of the gender difference in lifetime rates of major depression: a theoretical model. *Archives of General Psychiatry, 57*(1), 21-27. https://doi.org/10.1001/archpsyc.57.1.21

Dahl, R. E., & Gunnar, M. R. (2009). Heightened stress responsiveness and emotional reactivity during pubertal maturation: implications for psychopathology. *Development and Psychopathology, 21*(1), 1-6. https://doi.org/10.1017/S0954579409000017

Davis, J. P., Dumas, T. M., Merrin, G. J., Espelage, D. L., Tan, K., Madden, D., & Hong, J. S. (2018). Examining the pathways between bully victimization, depression, academic

achievement, and problematic drinking in adolescence. *Psychology of Addictive Behaviors, 32*(6), 605-616. https://doi.org/10.1037/adb0000394

Davis, M. C., Matthews, K. A., & Twamley, E. W. (1999). Is life more difficult on Mars or Venus? A meta-analytic review of sex differences in major and minor life events. *Annals of Behavioral Medicine, 21*(1), 83-97. https://doi.org/10.1007/BF02895038

De Bolle, M., Beyers, W., De Clercq, B., & De Fruyt, F. (2012). General personality and psychopathology in referred and nonreferred children and adolescents: an investigation of continuity, pathoplasty, and complication models. *Journal of Abnormal Psychology, 121*(4), 958-970. https://doi.org/10.1037/a0027742

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin, 141*(4), 858-900. https://doi.org/10.1037/a0038498

Dedovic, K., Engert, V., Duchesne, A., Lue, S. D., Andrews, J., Efanov, S. I., Beaudry, T., & Pruessner, J. C. (2010). Cortisol awakening response and hippocampal volume: vulnerability for major depressive disorder? *Biological Psychiatry, 68*(9), 847-853. https://doi.org/10.1016/j.biopsych.2010.07.025

DeNeve, K. M., & Cooper, H. (1998). The happy personality: a meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin, 124*(2), 197-229. https://doi.org/10.1037/0033-2909.124.2.197

Díaz, K. I., & Fite, P. J. (2019). Cyber victimization and its association with substance use, anxiety, and depression symptoms among middle school youth. *Child & Youth Care Forum, 48*, 529-544. https://doi.org/10.1007/s10566-019-09493-w

Doom, J. R., & Gunnar, M. R. (2013). Stress physiology and developmental psychopathology: past, present, and future. *Development and Psychopathology, 25*(4pt2), 1359-1373. https://doi.org/10.1017/S0954579413000667

Dreger, A. D., & Herndon, A. M. (2009). Progress and politics in the intersex rights movement: Feminist theory in action. *GLQ: A Journal of Lesbian and Gay Studies, 15*(2), 199-224. https://doi.org/10.1215/10642684-2008-134

Duncan, L., Comeau, J., Wang, L., Vitoroulis, I., Boyle, M. H., & Bennett, K. (2019). Research Review: Test–retest reliability of standardized diagnostic interviews to assess child and adolescent psychiatric disorders: a systematic review and meta-analysis. *Journal of Child Psychology and Psychiatry, 60*(1), 16-29. https://doi.org/10.1111/jcpp.12876

Dunn, C., & Sicouri, G. (2022). The relationship between loneliness and depressive symptoms in children and adolescents: A meta-analysis. *Behaviour Change, 39*(3), 134-145. https://doi.org/10.1017/bec.2022.13

Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology. *Journal of Child Psychology and*

Ejova, A., Milojev, P., Worthington Jr, E. L., Bulbulia, J., & Sibley, C. G. (2020). The Big Six personality traits and mental distress: Dynamic modeling in a population panel study reveals bidirectional relationships involving Neuroticism, Extraversion, and Conscientiousness. *Personality and Social Psychology Bulletin, 46*(9), 1287-1302. https://doi.org/10.1177/0146167219895349

ElHassan, N. O., Hall, R. W., Thomas, B. R., Palmer, T. W., Kaiser, J. R., & Li, C. (2023). Anxiety, depression, and behavioral and/or conduct disorder in adolescence among former preterm and term infants of different race and ethnicities. *Journal of Racial and Ethnic Health Disparities, 10*(3), 1379-1391. https://doi.org/10.1177/0146167219895349

Eng, A. G., Phan, J. M., Shirtcliff, E. A., Eisenlohr-Moul, T. A., Goh, P. K., & Martel, M. M. (2023). Aging and pubertal development differentially predict symptoms of ADHD, depression, and impairment in children and adolescents: An eight-year longitudinal study. *Research on Child and Adolescent Psychopathology*, 1-14. https://doi.org/10.1007/s10802-023-01030-7

Enright, R. D., Shukla, D. G., & Lapsley, D. K. (1980). Adolescent egocentrism-sociocentrism and self-consciousness. *Journal of Youth and Adolescence, 9*, 101-116. https://doi.org/10.1007/BF02087929

Essau, C., & Ollendick, T. (2009). *Diagnosis and Assessment of Adolescent Depression*. In S.N. Hoeksema & L.M. Hilt (Eds.). Handbook of Depression in Adolescents (pp. 33-47). Routledge/Taylor & Francis Group.

Felton, J. W., Cole, D. A., Havewala, M., Kurdziel, G., & Brown, V. (2019). Talking together, thinking alone: Relations among co-rumination, peer relationships, and rumination. *Journal of Youth and Adolescence, 48*, 731-743. https://doi.org/10.1007/s10964-018-0937-z

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.

Finsaas, M. C., Bufferd, S. J., Dougherty, L. R., Carlson, G. A., & Klein, D. N. (2018). Preschool psychiatric disorders: homotypic and heterotypic continuity through middle

childhood and early adolescence. *Psychological Medicine, 48*(13), 2159-2168. https://doi.org/10.1017/S0033291717003646

Fleming, K. A., Heintzelman, S. J., & Bartholow, B. D. (2016). Specifying associations between conscientiousness and executive functioning: Mental set shifting, not prepotent response inhibition or working memory updating. *Journal of Personality, 84*(3), 348-360. https://doi.org/10.1111/jopy.12163

Forbes, E. E., & Dahl, R. E. (2010). Pubertal development and behavior: hormonal activation of social and motivational tendencies. *Brain and Cognition, 72*(1), 66-72. https://doi.org/10.1016/j.bandc.2009.10.007

Ford, T., Goodman, R., & Meltzer, H. (2003). The British child and adolescent mental health survey 1999: the prevalence of DSM-IV disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*(10), 1203-1211. https://doi.org/10.1097/00004583-200310000-00011

Frankenberger, K. D. (2000). Adolescent egocentrism: A comparison among adolescents and adults. *Journal of Adolescence, 23*(3), 343-354. https://doi.org/10.1006/jado.2000.0319

Fraticelli, S., Caratelli, G., De Berardis, D., Ducci, G., Pettorruso, M., Martinotti, G., . . . Di Giannantonio, M. (2022). Gender differences in attention deficit hyperactivity disorder: An update of the current evidence. *Rivista di Psichiatria, 57*(4), 159-164. https://doi.org/10.1708/3855.38380

Fresson, M., Meulemans, T., Dardenne, B., & Geurten, M. (2019). Overdiagnosis of ADHD in boys: Stereotype impact on neuropsychological assessment. *Applied Neuropsychology: Child, 8*(3), 231-245. https://doi.org/10.1080/21622965.2018.1430576

Garber, J., Keiley, M. K., & Martin, N. C. (2002). Developmental trajectories of adolescents' depressive symptoms: predictors of change. *Journal of Consulting and Clinical Psychology, 70*(1), 79-95. https://doi.org/10.1037//0022-006x.70.1.79

Gau, S. S., Chong, M.-Y., Chen, T. H., & Cheng, A. T. (2005). A 3-year panel study of mental disorders among adolescents in Taiwan. *American Journal of Psychiatry, 162*(7), 1344-1350. https://doi.org/10.1176/appi.ajp.162.7.1344

Ge, X., Conger, R. D., & Elder Jr, G. H. (2001). Pubertal transition, stressful life events, and the emergence of gender differences in adolescent depressive symptoms. *Developmental Psychology, 37*(3), 404-417. https://doi.org/10.1037//0012-1649.37.3.404

Ge, X., Lorenz, F. O., Conger, R. D., Elder, G. H., & Simons, R. L. (1994). Trajectories of stressful life events and depressive symptoms during adolescence. *Developmental Psychology, 30*(4), 467-438. https://doi.org/10.1037/0012-1649.30.4.467

Gerhard, T. (2008). Bias: considerations for research practice. *American Journal of Health-System Pharmacy, 65*(22), 2159-2168. https://doi.org/10.2146/ajhp070369

Gibb, B. E. (2014). Depression in children. In I. H. Gotlib & C. Hammen, L. (Eds.), *Handbook of depression* (3rd ed., pp. 374-390). New York, NY: The Guilford Press.

Glickman, E. A., Choi, K. W., Lussier, A. A., Smith, B. J., & Dunn, E. C. (2021). Childhood emotional neglect and adolescent depression: assessing the protective role of peer social support in a longitudinal birth cohort. *Frontiers in Psychiatry, 12*, 1-9. https://doi.org/10.3389/fpsyt.2021.681176

Global Burden of Disease, M. D. C. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Psychiatry, 9*(12), 137-150. https://doi.org/10.1016/S2215-0366(21)00395-3

Goldstein, B. L., Kotov, R., Perlman, G., Watson, D., & Klein, D. N. (2018). Trait and facet-level predictors of first-onset depressive and anxiety disorders in a community sample of adolescent girls. *Psychological Medicine, 48*(8), 1282-1290. https://doi.org/10.1017/S0033291717002719

Goldstein, B. L., Perlman, G., Eaton, N. R., Kotov, R., & Klein, D. N. (2020). Testing explanatory models of the interplay between depression, neuroticism, and stressful life events: a dynamic trait-stress generation approach. *Psychological Medicine, 50*(16), 2780-2789. https://doi.org/10.1017/S0033291719002927

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry, 177*(6), 534-539. https://doi.org/10.1192/bjp.177.6.534

Grabe, S., Hyde, J. S., & Lindberg, S. M. (2007). Body objectification and depression in adolescents: The role of gender, shame, and rumination. *Psychology of Women Quarterly, 31*(2), 164-175. https://doi.org/10.1111/j.1471-6402.2007.00350.x

Graber, J. A. (2013). Pubertal timing and the development of psychopathology in adolescence and beyond. *Hormones and Behavior, 64*(2), 262-269. https://doi.org/10.1016/j.yhbeh.2013.04.003

Grant, K. E., Compas, B. E., Thurm, A. E., McMahon, S. D., & Gipson, P. Y. (2004). Stressors and child and adolescent psychopathology: Measurement issues and prospective effects. *Journal of Clinical Child and Adolescent Psychology, 33*(2), 412-425. https://doi.org/10.1207/s15374424jccp3302_23

Gray, S., & Daraganova, G. (2017). *Adolescent help-seeking. LSAC Annual Statistical Report* (Report No. 8). Australian Institute of Family Studies. https://growingupinaustralia.gov.au/research-findings/annual-statistical-report-2017/adolescent-help-seeking?_ga=2.150713526.1655597972.1693114289-86141104.1693114289

Guilliams, T. G., & Edwards, L. (2010). Chronic stress and the HPA axis. *The Standard, 9*(2), 1-12. https://www.pointinstitute.org/wp-content/uploads/2012/10/standard_v_9.2_hpa_axis.pdf

Gustavson, K., von Soest, T., Karevold, E., & Røysamb, E. (2012). Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study. *BMC Public Health, 12*, 1-11. https://doi.org/10.1186/1471-2458-12-918

Hagenauer, M. H., & Lee, T. M. (2012). The neuroendocrine control of the circadian system: adolescent chronotype. *Frontiers in Neuroendocrinology, 33*(3), 211-229. https://doi.org/10.1016/j.yfrne.2012.04.003

Hager, A. D., & Leadbeater, B. J. (2016). The longitudinal effects of peer victimization on physical health from adolescence to young adulthood. *Journal of Adolescent Health, 58*(3), 330-336. https://doi.org/10.1016/j.jadohealth.2015.10.014

Hakulinen, C., Elovainio, M., Pulkki-Råback, L., Böckerman, P., Viinikainen, J., Pehkonen, J., Raitakari, O. T., Keltikangas-Järvinen, L., & Hintsanen, M. (2016). Depressive symptoms and long-term income: The Young Finns Study. *Journal of Affective Disorders, 204*, 120-123. https://doi.org/10.1016/j.jad.2016.06.028

Hakulinen, C., Elovainio, M., Pulkki-Råback, L., Virtanen, M., Kivimäki, M., & Jokela, M. (2015). Personality and depressive symptoms: Individual participant meta-analysis of 10 cohort studies. *Depression and Anxiety, 32*(7), 461-470. https://doi.org/10.1002/da.22376

Halse, M., Steinsbekk, S., Hammar, Å., & Wichstrøm, L. (2022). Longitudinal relations between impaired executive function and symptoms of psychiatric disorders in childhood. *Journal of Child Psychology and Psychiatry, 63*(12), 1574-1582. https://doi.org/10.1111/jcpp.13622

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102-116. https://doi.org/10.1037/a0038889

Hamaker, E. L., Mulder, J. D., & van IJzendoorn, M. H. (2020). Description, prediction and causation: Methodological challenges of studying child and adolescent development. *Developmental Cognitive Neuroscience*, 100867, 1-14. https://doi.org/10.1016/j.dcn.2020.100867

Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology, 1*, 293-319. https://doi.org/10.1146/annurev.clinpsy.1.102803.143938

Hammen, C. (2009a). Adolescent depression: Stressful interpersonal contexts and risk for recurrence. *Current Directions in Psychological Science, 18*(4), 200-204. https://doi.org/10.1111/j.1467-8721.2009.01636.x

Hammen, C. (2009b). Stress exposure and stress generation in adolescent depression. In S. Nolen-Hoeksema & L. M. Hilt (Eds.), *Handbook of depression in adolescents* (pp. 305-333). Routledge/Taylor & Francis Group.

Hampel, P., & Petermann, F. (2005). Age and gender effects on coping in children and adolescents. *Journal of Youth and Adolescence, 34*(2), 73-83. https://doi.org/10.1007/s10964-005-3207-9

Hamza, C. A., & Willoughby, T. (2011). Perceived parental monitoring, adolescent disclosure, and adolescent depressive symptoms: A longitudinal examination. *Journal of Youth and Adolescence, 40*, 902-915. https://doi.org/10.1007/s10964-010-9604-8

Hancock, J., Liu, S. X., Luo, M., & Mieczkowski, H. (2022). Psychological well-being and social media use: A meta-analysis of associations between social media use and depression, anxiety, loneliness, eudaimonic, hedonic and social well-being. *Anxiety, Loneliness, Eudaimonic, Hedonic and Social Well-Being*, 1-62. https://doi.org/10.2139/ssrn.4053961

Hankin, B. L. (2008). Cognitive vulnerability–stress model of depression during adolescence: Investigating depressive symptom specificity in a multi-wave prospective study. *Journal of Abnormal Child Psychology, 36*(7), 999-1014. https://doi.org/10.1007/s10802.008.9228.6

Hankin, B. L., & Abramson, L. Y. (2001). Development of gender differences in depression: An elaborated cognitive vulnerability–transactional stress theory. *Psychological Bulletin, 127*(6), 773-796. https://doi.org/10.1037/0033-2909.127.6.773

Hankin, B. L., Mermelstein, R., & Roesch, L. (2007). Sex differences in adolescent depression: Stress exposure and reactivity models. *Child Development, 78*(1), 279-295. https://doi.org/10.1111/j.1467-8624.2007.00997.x

Hankin, B. L., Stone, L., & Wright, P. A. (2010). Corumination, interpersonal stress generation, and internalizing symptoms: Accumulating effects and transactional influences in a multiwave study of adolescents. *Development and Psychopathology, 22*(1), 217-235. https://doi.org/10.1017/S0954579409990368

Hankin, B. L., Young, J. F., Abela, J. R., Smolen, A., Jenness, J. L., Gulley, L. D., Technow, J. R., Gottlieb, A. B., Cohen, J. R., & Oppenheimer, C. W. (2015). Depression from childhood into late adolescence: Influence of gender, development, genetic susceptibility, and peer stress. *Journal of Abnormal Psychology, 124*(4), 803-816. https://doi.org/10.1037/abn0000089

Hare, T. A., Tottenham, N., Galvan, A., Voss, H. U., Glover, G. H., & Casey, B. J. (2008). Biological substrates of emotional reactivity and regulation in adolescence during an emotional go-nogo task. *Biological Psychiatry, 63*(10), 927-934. https://doi.org/10.1016/j.biopsych.2008.03.015

Harter, S. (2006). The self. In N. Eisenberg, W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 505–570). John Wiley & Sons.

Harter, S. (2015). *The construction of the self: Developmental and sociocultural foundations*. Guilford Publications.

Harter, S., Waters, P., & Whitesell, N. R. (1998). Relational self-worth: Differences in perceived worth as a person across interpersonal contexts among adolescents. *Child Development, 69*(3), 756-766. https://doi.org/10.1111/j.1467-8624.1998.tb06241.x

Hemphill, S. A., Tollit, M., Kotevski, A., & Heerde, J. A. (2015). Predictors of traditional and cyber-bullying victimization: A longitudinal study of Australian secondary school students. *Journal of Interpersonal Violence*, *30*(15), 2567-2590. https://doi.org/10.1177/0886260514553636

Hoferichter, F., Raufelder, D., & Eid, M. (2014). The mediating role of socio-motivational relationships in the interplay of perceived stress, neuroticism, and test anxiety among adolescent students. *Psychology in the Schools, 51*(7), 736-752. https://doi.org/10.1002/pits.21778

Holfeld, B., & Leadbeater, B. J. (2017). Concurrent and longitudinal associations between early adolescents' experiences of school climate and cyber victimization. *Computers in Human Behavior, 76*, 321-328. https://doi.org/10.1016/j.chb.2017.07.037

Hu, Y., Bai, Y., Pan, Y., & Li, S. (2021). Cyberbullying victimization and depression among adolescents: A meta-analysis. *Psychiatry Research, 305*, 1-8. https://doi.org/10.1016/j.psychres.2021.114198

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist, 74*(2), 171-193. https://doi.org/10.1037/amp0000307

Hyde, J. S., & Mezulis, A. H. (2020). Gender differences in depression: biological, affective, cognitive, and sociocultural factors. *Harvard Review of Psychiatry, 28*(1), 4-13. https://doi.org/10.1097/HRP.0000000000000230

Hyde, J. S., Mezulis, A. H., & Abramson, L. Y. (2008). The ABCs of depression: integrating affective, biological, and cognitive models to explain the emergence of the gender difference in depression. *Psychological Review, 115*(2), 291-313. https://doi.org/10.1037/0033-295X.115.2.291

Hysing, M., Pallesen, S., Stormark, K. M., Jakobsen, R., Lundervold, A. J., & Sivertsen, B. (2015). Sleep and use of electronic devices in adolescence: results from a large population-based study. *BMJ Open, 5*(1), 1-7. http://doi.org/10.1136/bmjopen-2014-006748

Ivie, E. J., Pettitt, A., Moses, L. J., & Allen, N. B. (2020). A meta-analysis of the association between adolescent social media use and depressive symptoms. *Journal of Affective Disorders, 275*, 165-174. https://doi.org/10.1016/j.jad.2020.06.014

Jadva, V., Guasp, A., Bradlow, J., Bower-Brown, S., & Foley, S. (2023). Predictors of self-harm and suicide in LGBT youth: The role of gender, socio-economic status, bullying and school experience. *Journal of Public Health, 45*(1), 102-108. https://doi.org/10.1093/pubmed/fdab383

Jenness, J. L., Peverill, M., King, K. M., Hankin, B. L., & McLaughlin, K. A. (2019). Dynamic associations between stressful life events and adolescent internalizing psychopathology in a multiwave longitudinal study. *Journal of Abnormal Psychology, 128*(6), 596-609. https://doi.org/10.1037/abn0000450

Jenni, O. G., Achermann, P., & Carskadon, M. A. (2005). Homeostatic sleep regulation in adolescents. *Sleep, 28*(11), 1446-1454. https://doi.org/10.1093/sleep/28.11.1446

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). The Guilford Press.

Johnson, D., Dupuis, G., Piche, J., Clayborne, Z., & Colman, I. (2018). Adult mental health outcomes of adolescent depression: A systematic review. *Depression and Anxiety, 35*(8), 700-716. https://doi.org/10.1002/da.22777

Joiner, T., Coyne, J. C., & Blalock, J. (1999). On the interpersonal nature of depression: Overview and synthesis. In T. Joiner & J. C. Coyne (Eds.), *The interactional nature of depression: Advances in interpersonal approaches* (pp. 3–19). American Psychological Association.

Kaufman, J., Birmaher, B., Axelson, D., Perepletchikova, F., Brent, D., & Ryan, N. (2016). K-SADS-PL DSM-5. *Western Psychiatric Institute and Clinic*.

Kelly, R. J., & El-Sheikh, M. (2014). Reciprocal relations between children's sleep and their adjustment over time. *Developmental Psychology, 50*(4), 1137-47. https://doi.org/10.1037/a0034501

Kendler, K. S., & First, M. B. (2010). Alternative futures for the DSM revision process: iteration v. paradigm shift. *The British Journal of Psychiatry, 197*(4), 263-265. http://doi.org/10.1192/bjp.bp.109.076794

Kendler, K. S., Gardner, C. O., Neale, M. C., Aggen, S., Heath, A., Colodro-Conde, L., Couvyduchesne, B., Byrne, E. M., Martin, N. G., & Gillespie, N. A. (2019). Shared and specific genetic risk factors for lifetime major depression, depressive symptoms and neuroticism in three population-based twin samples. *Psychological Medicine, 49*(16), 2745-2753. https://doi.org/10.1017/S003329171800377X

Kercher, A. J., Rapee, R. M., & Schniering, C. A. (2009). Neuroticism, life events and negative thoughts in the development of depression in adolescent girls. *Journal of Abnormal Child Psychology, 37*, 903-915. https://doi.org/10.1007/s10802-009-9325-1

Kessler, R. C., & Üstün, T. B. (2004). The world mental health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). *International Journal of Methods in Psychiatric Research, 13*(2), 93-121. https://doi.org/10.1002/mpr.168

Khoo, S., & Simms, L. J. (2018). Links between depression and openness and its facets. *Personality and Mental Health, 12*(3), 203-215. https://doi.org/10.1002/pmh.1417

Klefsjö, U., Kantzer, A. K., Gillberg, C., & Billstedt, E. (2021). The road to diagnosis and treatment in girls and boys with ADHD–gender differences in the diagnostic process. *Nordic Journal of Psychiatry, 75*(4), 301-305. https://doi.org/10.1080/08039488.2020.1850859

Klein, D. N. (2023). Assessment of Depression in Adults and Youth. *Assessment*, 1-16. https://doi.org/10.1177/10731911231167446

Klein, D. N., Kotov, R., & Bufferd, S. J. (2011). Personality and depression: explanatory models and review of the evidence. *Annual Review of Clinical Psychology, 7*, 269-295. https://doi.org/10.1146/annurev-clinpsy-032210-104540

Klein, D. N., Shankman, S. A., & Rose, S. (2006). Ten-year prospective follow-up study of the naturalistic course of dysthymic disorder and double depression. *American Journal of Psychiatry, 163*(5), 872-880. https://doi.org/10.1176/ajp.2006.163.5.872

Klimstra, T. A., Akse, J., Hale III, W. W., Raaijmakers, Q. A., & Meeus, W. H. (2010). Longitudinal associations between personality traits and problem behavior symptoms in adolescence. *Journal of Research in Personality, 44*(2), 273-284. https://doi.org/10.1016/j.jrp.2010.02.004

Klinger, D. A., Freeman, J. G., Bilz, L., Liiv, K., Ramelow, D., Sebok, S. S., Semdal, U., Dür, W., & Rasmussen, M. (2015). Cross-national trends in perceived school pressure by gender and age from 1994 to 2010. *The European Journal of Public Health, 25*(suppl_2), 51-56. https://doi.org/10.1093/eurpub/ckv027

Klomek, A. B., Barzilay, S., Apter, A., Carli, V., Hoven, C. W., Sarchiapone, M., Hadlaczky, G., Balazs, J. Kereszteny, A., Brunner, R., Kaess, M., Bobes, J., Saiz, P. A., Cosman, D., Haring, C., Banzer, R., McMahon, E., Keeley, H., Kahn, J.-P., . . . Wassermann, D. (2019). Bi-directional longitudinal associations between different types of bullying victimization, suicide ideation/attempts, and depression among a large sample of European adolescents. *Journal of Child Psychology and Psychiatry, 60*(2), 209-215. https://doi.org/10.1111/jcpp.12951

Kochel, K. P., & Rafferty, D. (2020). Prospective associations between children's depressive symptoms and peer victimization: The role of social helplessness. *British Journal of Developmental Psychology, 38*(1), 15-30. https://doi.org/10.1111/bjdp.12301

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Kovacs, M., Obrosky, S., & George, C. (2016). The course of major depressive disorder from childhood to young adulthood: Recovery and recurrence in a longitudinal observational study. *Journal of AffectiveDdisorders, 203*, 374-381. https://doi.org/10.1016/j.jad.2016.05.042

Krause, E. D., Vélez, C. E., Woo, R., Hoffmann, B., Freres, D. R., Abenavoli, R. M., & Gillham, J. E. (2018). Rumination, depression, and gender in early adolescence: A longitudinal study of a bidirectional model. *The Journal of Early Adolescence, 38*(7), 923-946. https://doi.org/10.1177/0272431617704

Krause, H.-V., Baum, K., Baumann, A., & Krasnova, H. (2021). Unifying the detrimental and beneficial effects of social network site use on self-esteem: a systematic literature review. *Media Psychology, 24*(1), 10-47. https://doi.org/10.1080/15213269.2019.1656646

Kuehner, C. (2017). Why is depression more common among women than among men? *The Lancet Psychiatry, 4*(2), 146-158. https://doi.org/10.1016/S2215-0366(16)30263-2

Lahey, B. B., Applegate, B., Waldman, I. D., Loft, J. D., Hankin, B. L., & Rick, J. (2004). The structure of child and adolescent psychopathology: generating new hypotheses. *Journal of Abnormal Psychology, 113*(3), 358-385. https://doi.org/10.1037/0021-843X.113.3.358

Larsen, B., & Luna, B. (2018). Adolescence as a neurobiological critical period for the development of higher-order cognition. *Neuroscience & Biobehavioral Reviews, 94*, 179-195. https://doi.org/10.1016/j.neubiorev.2018.09.005

Larson, R., & Ham, M. (1993). Stress and" storm and stress" in early adolescence: The relationship of negative events with dysphoric affect. *Developmental Psychology, 29*(1), 130-140. https://doi.org/10.1037/0012-1649.29.1.130

Lavigne, J. V., LeBailly, S. A., Hopkins, J., Gouze, K. R., & Binns, H. J. (2009). The prevalence of ADHD, ODD, depression, and anxiety in a community sample of 4-year-olds. *Journal of Clinical Child & Adolescent Psychology, 38*(3), 315-328. https://doi.org/10.1080/15374410902851382

Le Couteur, A., & Gardner, F. (2008). Use of structured interviews and observational methods in clinical settings. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor & A. Thapar (Eds.), *Rutter's child and adolescent psychiatry* (pp. 271–288). Wiley Blackwell.

Lepore, S. J., & Kliewer, W. (2019). Social intelligence attenuates association between peer victimization and depressive symptoms among adolescents. *Psychology of Violence, 9*(6), 644-652. https://doi.org/10.1037/vio0000234

Li, J., Yao, G., Liu, S., Li, X., Zhao, W., Du, X., Wang, X., Sun, J., & Xu, Y. (2021). Mechanisms of the effects of parental emotional warmth on extraversion in children and adolescents. *Neuroscience, 467*, 134-141. https://doi.org/10.1016/j.neuroscience.2021.05.021

Li, Y., Gu, S., Wang, Z., Li, H., Xu, X., Zhu, H., Deng, S., Ma, X., Feng, G., Wang, F., & Huang, J. H. (2019). Relationship between stressful life events and sleep quality: rumination as a mediator and resilience as a moderator. *Frontiers in Psychiatry, 10*(348), 1-9. https://doi.org/10.3389/fpsyt.2019.00348

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198-1202. https://doi.org/10.1080/01621459.1988.10478722

Liu, C., Chen, L., & Chen, S. (2020). Influence of neuroticism on depressive symptoms among Chinese adolescents: the mediation effects of cognitive emotion regulation strategies. *Frontiers in Psychiatry, 11*(420), 1-9. https://doi.org/10.3389/fpsyt.2020.00420

Liu, H. (2022). *The Influence of Parental Divorce on Depression Symptoms in Adolescence and Early Adulthood: A Systematic Review and Meta-Analysis.* [Paper presentation]. The 2021 International Conference on Public Art and Human Development (ICPAHD 2021). https://doi.org/10.2991/assehr.k.220110.092

Liu, M., Kamper-DeMarco, K. E., Zhang, J., Xiao, J., Dong, D., & Xue, P. (2022). Time Spent on Social Media and Risk of Depression in Adolescents: A Dose–Response Meta-Analysis. *International Journal of Environmental Research and Public health, 19*(9), 1-17. https://doi.org/10.3390/ijerph19095164

Lovato, N., & Gradisar, M. (2014). A meta-analysis and model of the relationship between sleep and depression in adolescents: recommendations for future research and clinical practice. *Sleep Medicine Reviews, 18*(6), 521-529. https://doi.org/10.1016/j.smrv.2014.03.006

Luby, J. L., Heffelfinger, A. K., Mrakotsky, C., Hessler, M. J., Brown, K. M., & Hildebrand, T. (2002). Preschool major depressive disorder: preliminary validation for developmentally modified DSM-IV criteria. *Journal of the American Academy of*

Child & Adolescent Psychiatry, 41(8), 928-937. https://doi.org/10.1097/00004583-200208000-00011

Ludwig-Walz, H., Dannheim, I., Pfadenhauer, L. M., Fegert, J. M., & Bujard, M. (2022). Increase of depression among children and adolescents after the onset of the COVID-19 pandemic in Europe: a systematic review and meta-analysis. *Child and Adolescent Psychiatry and Mental Health, 16*(1), 1-20. https://doi.org/10.1186/s13034-022-00546-y

Madsen, A., Oehme, N. B., Roelants, M., Bruserud, I. S., Eide, G. E., Viste, K. Bjerknes, R., Almås, B., Rosendahl, K., Sagen, J. V., Mellgren, G., Juliusson, P. R. (2020). Testicular ultrasound to stratify hormone references in a cross-sectional Norwegian study of male puberty. *The Journal of Clinical Endocrinology & Metabolism, 105*(6), 1888-1898. https://doi.org/10.1210/clinem/dgz094

Markkula, N., Marola, N., Nieminen, T., Koskinen, S., Saarni, S. I., Härkänen, T., & Suvisaari, J. (2017). Predictors of new-onset depressive disorders–Results from the longitudinal Finnish Health 2011 Study. *Journal of Affective Disorders, 208*, 255-264. https://doi.org/10.1016/j.jad.2016.08.051

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107-123. https://doi.org/10.1207/s15326985ep2003_1

Martínez-San Miguel, Y., & Tobias, S. (Eds.) (2016). *Trans studies: The challenge to hetero/homo normativities*. Rutgers University Press.

Mason, W. A., Chmelka, M. B., Trudeau, L., & Spoth, R. L. (2017). Gender moderation of the intergenerational transmission and stability of depressive symptoms from early adolescence to early adulthood. *Journal of Youth and Adolescence, 46*(1), 248-260. https://doi.org/10.1007/s10964-016-0480-8

Masselink, M., Van Roekel, E., Hankin, B., Keijsers, L., Lodder, G., Vanhalst, J., Verhagen, M., Young, J. F., & Oldehinkel, A. (2018). The longitudinal association between self–esteem and depressive symptoms in adolescents: Separating between–person effects from within–person effects. *European Journal of Personality, 32*(6), 653-671. https://doi.org/10.1002/per.2179

Mathew, G. M., Hale, L., & Chang, A.-M. (2019). Sex moderates relationships among school night sleep duration, social jetlag, and depressive symptoms in adolescents. *Journal of Biological Rhythms, 34*(2), 205-217. https://doi.org/10.1177/0748730419828102

Mayes, S. D., Calhoun, S. L., Waxmonsky, J. G., Kokotovich, C., Baweja, R., Lockridge, R., & Bixler, E. O. (2019). Demographic differences in disruptive mood dysregulation disorder symptoms in ADHD, autism, and general population samples. *Journal of Attention Disorders, 23*(8), 849-858. https://doi.org/10.1177/1087054716664409

McCrae, R. R., & Costa Jr, P. T. (2008). The five-factor theory of personality. In B. Gregory, D. H. Saklofske, & Matthews, G. (Eds), *The Sage Handbook of Personality Theory and Assessment* (pp. 273-294). Sage Publications.

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

McLaughlin, K. A., Hatzenbuehler, M. L., & Hilt, L. M. (2009). Emotion dysregulation as a mechanism linking peer victimization to internalizing symptoms in adolescents. *Journal of Consulting and Clinical Psychology, 77*(5), 894-904. https://doi.org/10.1037/a0015760

McLean, S. A., Jarman, H. K., & Rodgers, R. F. (2019). How do "selfies" impact adolescents' well-being and body confidence? A narrative review. *Psychology Research and Behavior Management, 12*, 513-521. https://doi.org/10.2147/PRBM.S177834

Meiran, N., Diamond, G. M., Toder, D., & Nemets, B. (2011). Cognitive rigidity in unipolar depression and obsessive compulsive disorder: Examination of task switching, Stroop, working memory updating and post-conflict adaptation. *Psychiatry Research, 185*(1-2), 149-156. https://doi.org/10.1016/j.psychres.2010.04.044

Mendle, J. (2014). Why puberty matters for psychopathology. *Child Development Perspectives, 8*(4), 218-222. https://doi.org/10.1111/cdep.12092

Mendle, J., Beam, C. R., McKone, K. M., & Koch, M. K. (2020). Puberty and transdiagnostic risks for mental health. *Journal of Research on Adolescence, 30*(3), 687-705. https://doi.org/10.1111/jora.12552

Mendle, J., & Ferrero, J. (2012). Detrimental psychological outcomes associated with pubertal timing in adolescent boys. *Developmental Review, 32*(1), 49-66. https://doi.org/10.1016/j.dr.2011.11.001

Merikangas, K. R., He, J.-P., Brody, D., Fisher, P. W., Bourdon, K., & Koretz, D. S. (2010a). Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. *Pediatrics, 125*(1), 75-81. https://doi.org/10.1542/peds.2008-2598

Merikangas, K. R., He, J.-P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010b). Lifetime prevalence of mental disorders in

US adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry, 49*(10), 980-989. https://doi.org/10.1016/j.jaac.2010.05.017

Mikolas, P., Tozzi, L., Doolin, K., Farrell, C., O'Keane, V., & Frodl, T. (2018). Hippocampal subfields volumes, endocrine stress axis and early life stress in major depressive disorder. *European Neuropsychopharmacology, 28*, 71-72. https://doi.org/10.1016/j.euroneuro.2017.12.101

Merriam-Webster (2023). *Medical definition of disease.* Retrieved 25th of July from http://www.merriam-webster.com/medical/disease

Moreno-Agostino, D., Wu, Y.-T., Daskalopoulou, C., Hasan, M. T., Huisman, M., & Prina, M. (2021). Global trends in the prevalence and incidence of depression: a systematic review and meta-analysis. *Journal of Affective Disorders, 281*, 235-243. https://doi.org/10.1016/j.jad.2020.12.035

Morssinkhof, M., Van Wylick, D., Priester-Vink, S., van der Werf, Y., den Heijer, M., van den Heuvel, O., & Broekman, B. (2020). Associations between sex hormones, sleep problems and depression: A systematic review. *Neuroscience & Biobehavioral Reviews, 118*, 669-680. https://doi.org/10.1016/j.neubiorev.2020.08.006

Motro, D., & Ellis, A. P. (2017). Boys, don't cry: Gender and reactions to negative performance feedback. *Journal of Applied Psychology, 102*(2), 227-235. https://doi.org/10.1037/apl0000175

Mu, W., Luo, J., Rieger, S., Trautwein, U., & Roberts, B. W. (2019). The relationship between self-esteem and depression when controlling for neuroticism. *Collabra: Psychology, 5*(1), 1-13. https://doi.org/10.1525/collabra.204

Mueller, A. E., & Segal, D. L. (2014). Structured versus semistructured versus unstructured interviews. *The Encyclopedia of Clinical Psychology*, 1-7. https://doi.org/10.1002/9781118625392.wbecp069

Mulder, J. D., & Hamaker, E. L. (2020). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(4), 638-648. https://doi.org/10.1080/10705511.2020.1784738

Mund, M., & Nestler, S. (2019). Beyond the cross-lagged panel model: Next-generation statistical tools for analyzing interdependencies across the life course. *Advances in Life Course Research, 41*, 1-11. https://doi.org/10.1016/j.alcr.2018.10.002

Naragon-Gainey, K., & Simms, L. J. (2017). Clarifying the links of conscientiousness with internalizing and externalizing psychopathology. *Journal of Personality, 85*(6), 880-892. https://doi.org/10.1111/jopy.12295

Negriff, S., Fung, M. T., & Trickett, P. K. (2008). Self-rated pubertal development, depressive symptoms and delinquency: Measurement issues and moderation by gender and maltreatment. *Journal of Youth and Adolescence, 37*, 736-746. https://doi.org/10.1007/s10964-008-9274-y

Negriff, S., & Susman, E. J. (2011). Pubertal timing, depression, and externalizing problems: A framework, review, and examination of gender differences. *Journal of Research on Adolescence, 21*(3), 717-746. https://doi.org/10.1111/j.1532-7795.2010.00708.x

Nigg, J. T. (2006). Temperament and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 47*(3-4), 395-422. https://doi.org/10.1111/j.1469-7610.2006.01612.x

Nolen-Hoeksema, S., & Girgus, J. S. (1994). The emergence of gender differences in depression during adolescence. *Psychological Bulletin, 115*(3), 424-443. https://doi.org/10.1037/0033-2909.115.3.424

Nolen-Hoeksema, S., & Hilt, L. M. (Eds.) (2013). *Handbook of depression in adolescents*. Routledge/Taylor & Francis Group.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science, 3*(5), 400-424. https://doi.org/10.1111/j.1745-6924.2008.00088.x

Oehme, N., Bruserud, I. S., Madsen, A., & Júlíusson, P., B. (2020). Starter puberteten tidligere enn før? *Tidsskrift for Norsk Legeforening*. https://doi.org/10.4045/tidsskr.20.0043

Oldehinkel, A. J., & Bouma, E. M. (2011). Sensitivity to the depressogenic effect of stress and HPA-axis reactivity in adolescence: a review of gender differences. *Neuroscience & Biobehavioral Reviews, 35*(8), 1757-1770. https://doi.org/10.1016/j.neubiorev.2010.10.013

Ormel, J., Jeronimus, B. F., Kotov, R., Riese, H., Bos, E. H., Hankin, B., Rosmalen, J. G. M., & Oldehinkel, A. J. (2013). Neuroticism and common mental disorders: Meaning and utility of a complex relationship. *Clinical Psychology Review, 33*(5), 686-697. https://doi.org/10.1016/j.cpr.2013.04.003

Parker, G., & Brotchie, H. (2010). Gender differences in depression. *International Review of Psychiatry, 22*(5), 429-436. https://doi.org/10.3109/09540261.2010.492391

Paxton, S. J., Neumark-Sztainer, D., Hannan, P. J., & Eisenberg, M. E. (2006). Body dissatisfaction prospectively predicts depressive mood and low self-esteem in adolescent girls and boys. *Journal of Clinical Child and Adolescent Psychology, 35*(4), 539-549. https://doi.org/10.1207/s15374424jccp3504_5

Peterson, A. C., & Taylor, B. (1980). The biological approach to adolescence: Biological change and psychological adaptation In J. Adelson (Ed.), *Handbook of Adolescent Psychology*. John, Wiley & sons, Inc.

Pettersen, K. T. (2013). A study of shame from sexual abuse within the context of a Norwegian incest center. *Journal of Child Sexual Abuse, 22*(6), 677-694. https://doi.org/10.1080/10538712.2013.811139

Pfeifer, J. H., & Allen, N. B. (2021). Puberty initiates cascading relationships between neurodevelopmental, social, and internalizing processes across adolescence. *Biological Psychiatry, 89*(2), 99-108. https://doi.org/10.1016/j.biopsych.2020.09.002

Platt, J. M., Bates, L., Jager, J., McLaughlin, K. A., & Keyes, K. M. (2021). Is the US gender gap in depression changing over time? A meta-regression. *American Journal of Epidemiology, 190*(7), 1190-1206. https://doi.org/10.1093/aje/kwab002

Rankin, J. L., Lane, D. J., Gibbons, F. X., & Gerrard, M. (2004). Adolescent self-consciousness: Longitudinal age changes and gender differences in two cohorts. *Journal of Research on Adolescence, 14*(1), 1-21. https://doi.org/10.1111/j.1532-7795.2004.01401001.x

Rantanen, J., Metsäpelto, R. L., Feldt, T., Pulkkinen, L., & Kokko, K. (2007). Long-term stability in the Big Five personality traits in adulthood. *Scandinavian Journal of Psychology, 48*(6), 511-518. https://doi.org/10.1111/j.1467-9450.2007.00609.x

Rapee, R. M., Oar, E. L., Johnco, C. J., Forbes, M. K., Fardouly, J., Magson, N. R., & Richardson, C. E. (2019). Adolescent development and risk for the onset of social-emotional disorders: A review and conceptual model. *Behaviour Research and Rherapy, 123*(103501), 1-14. https://doi.org/10.1016/j.brat.2019.103501

Reardon, L. E., Leen-Feldner, E. W., & Hayward, C. (2009). A critical review of the empirical literature on the relation between anxiety and puberty. *Clinical Psychology Review, 29*(1), 1-23. https://doi.org/10.1016/j.cpr.2008.09.005

Richards, D. (2011). Prevalence and clinical course of depression: a review. *Clinical Psychology Review, 31*(7), 1117-1125. https://doi.org/10.1016/j.cpr.2011.07.004

Rickwood, D., Deane, F. P., Wilson, C. J., & Ciarrochi, J. (2005). Young people's help-seeking for mental health problems. *Australian E-journal for the Advancement of Mental Health, 4*(3), 218-251. https://doi.org/10.5172/jamh.4.3.218

Rimes, K. A., Goodship, N., Ussher, G., Baker, D., & West, E. (2019). Non-binary and binary transgender youth: Comparison of mental health, self-harm, suicidality, substance use and victimization experiences. *International Journal of Transgenderism, 20*(2-3), 230-240. https://doi.org/10.1080/15532739.2017.1370627

Roberts, A. G., & Lopez-Duran, N. L. (2019). Developmental influences on stress response systems: Implications for psychopathology vulnerability in adolescence. *Comprehensive Psychiatry, 88*, 9-21. https://doi.org/10.1016/j.comppsych.2018.10.008

Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Global self-esteem across the life span. *Psychology and Aging, 17*(3), 423-434. https://doi.org/10.1037/0882-7974.17.3.423

Rohde, P., Lewinsohn, P. M., & Seeley, J. R. (1994). Are adolescents changed by an episode of major depression? *Journal of the American Academy of Child & Adolescent Psychiatry, 33*(9), 1289-1298. https://doi.org/10.1097/00004583-199411000-00010

Romans, S. E., Tyas, J., Cohen, M. M., & Silverstone, T. (2007). Gender differences in the symptoms of major depressive disorder. *The Journal of Nervous and Mental Disease, 195*(11), 905-911. https://doi.org/10.1097/NMD.0b013e3181594cb7

Ronold, E. H., Schmid, M. T., Oedegaard, K. J., & Hammar, Å. (2020). A longitudinal 5-year follow-up study of cognitive function after first episode major depressive disorder: exploring state, scar and trait effects. *Frontiers in Psychiatry, 11*, 575867. 1-12. https://doi.org/10.3389/fpsyt.2020.575867

Rose, A. J. (2002). Co–rumination in the friendships of girls and boys. *Child Development, 73*(6), 1830-1843. https://doi.org/10.1111/1467-8624.00509

Rose, A. J. (2021). The costs and benefits of co-rumination. *Child Development Perspectives, 15*(3), 176-181. https://doi.org/10.1111/cdep.12419

Rose, A. J., Glick, G. C., Smith, R. L., Schwartz-Mette, R. A., & Borowski, S. K. (2017). Co-rumination exacerbates stress generation among adolescents with depressive symptoms. *Journal of Abnormal Child Psychology, 45*, 985-995. https://doi.org/10.1007/s10802-016-0205-1

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: potential trade-offs for the emotional and behavioral development of girls

and boys. *Psychological Bulletin, 132*(1), 98-131. https://doi.org/10.1037/0033-2909.132.1.98

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review*, 141-156. https://doi.org/10.2307/2096350

Rothbart, M. K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science, 16*(4), 207-212. https://doi.org/10.1111/j.1467-8721.2007.00505.x

Rudolph, K. D. (2009). The interpersonal context of adolescent depression. In S. Nolen-Hoeksema & L. M. Hilt (Eds.), *Handbook of depression in adolescents* (pp. 337-418). Routledge/Taylor & Francis Group.

Rudolph, K. D., & Flynn, M. (2014). Depression in adolescents. In I. H. Gotlb & C. Hammen, L (Eds.), *Handbook of depression* (3rd ed., pp. 391-409). The Guilford Press.

Rudolph, K. D., & Klein, D. N. (2009). Exploring depressive personality traits in youth: origins, correlates, and developmental consequences. *Developmental Psychopathology, 21*(4), 1155-1180. https://doi.org/10.1017/s0954579409990095

Rueger, S. Y., Malecki, C. K., Pyun, Y., Aycock, C., & Coyle, S. (2016). A meta-analytic review of the association between perceived social support and depression in childhood and adolescence. *Psychological Bulletin, 142*(10), 1017-1067. https://doi.org/10.1037/bul0000058

Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: meta-analyses of diagnoses and symptoms. *Psychological Bulletin, 143*(8), 783-822. https://doi.org/10.1037/bul0000102

Sandstrom, A., Perroud, N., Alda, M., Uher, R., & Pavlova, B. (2021). Prevalence of attention-deficit/hyperactivity disorder in people with mood disorders: A systematic review and meta-analysis. *Acta Psychiatrica Scandinavica, 143*(5), 380-391. https://doi.org/10.1111/acps.13283

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514. https://doi.org/10.1007/BF02296192

Satterwhite, A. K., & Luchner, A. F. (2016). Exploring the Relationship among Perceived Resilience, Dependency, and Self-Criticism: The Role of Culture and Social Support. *North American Journal of Psychology, 18*(1), 71-84. https://www.researchgate.net/publication/303645681

Scherr, S., Toma, C. L., & Schuster, B. (2018). Depression as a predictor of Facebook surveillance and envy. *Journal of Media Psychology, 31*(4), 196-202. https://doi.org/10.1027/1864-1105/a000247

Schmid, M., & Hammar, Å. (2013). A follow-up study of first episode major depressive disorder. Impairment in inhibition and semantic fluency—potential predictors for relapse? *Frontiers in Psychology, 4*(633), 1-13. https://doi.org/10.3389/fpsyg.2013.00633

Schramm, E., Klein, D. N., Elsaesser, M., Furukawa, T. A., & Domschke, K. (2020). Review of dysthymia and persistent depressive disorder: history, correlates, and clinical implications. *The Lancet Psychiatry, 7*(9), 801-812. https://doi.org/10.1016/S2215-0366(20)30099-7

Schulterbrandt, J. G., & Raskin, A. (1977). *Depression in childhood: Diagnosis, treatment, and conceptual models*. Raven Press (ID).

Segrin, C., & Dillard, J. P. (1992). The interactional theory of depression: A meta-analysis of the research literature. *Journal of Social and Clinical Psychology, 11*(1), 43-70. https://doi.org/10.1521/jscp.1992.11.1.43

Sharma, V., Coleman, S., Nixon, J., Sharples, L., Hamilton-Shield, J., Rutter, H., & Bryant, M. (2019). A systematic review and meta-analysis estimating the population prevalence of comorbidities in children and adolescents aged 5 to 18 years. *Obesity Reviews, 20*(10), 1341-1349. https://doi.org/10.1111/obr.12904

Shaw, T., Campbell, M. A., Eastham, J., Runions, K. C., Salmivalli, C., & Cross, D. (2019). Telling an adult at school about bullying: Subsequent victimization and internalizing problems. *Journal of Child and Family Studies, 28*, 2594-2605. https://doi.org/10.1007/s10826-019-01507-4

Shean, G., & Baldwin, G. (2008). Sensitivity and specificity of depression questionnaires in a college-age sample. *The Journal of Genetic Psychology, 169*(3), 281-292. https://doi.org/10.3200/GNTP.169.3.281-292

Shih, J. H., Eberhart, N. K., Hammen, C. L., & Brennan, P. A. (2006). Differential exposure and reactivity to interpersonal stress predict sex differences in adolescent depression. *Journal of Clinical Child and Adolescent Psychology, 35*(1), 103-115. https://doi.org/10.1207/s15374424jccp3501_9

Shiner, R., & Caspi, A. (2003). Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of Child Psychology and Psychiatry, 44*(1), 2-32. https://doi.org/10.1111/1469-7610.00101

Shorey, S., Ng, E. D., & Wong, C. H. (2022). Global prevalence of depression and elevated depressive symptoms among adolescents: A systematic review and meta-analysis. *British Journal of Clinical Psychology, 61*(2), 287-305. https://doi.org/10.1111/bjc.12333

Short, M. A., Booth, S. A., Omar, O., Ostlundh, L., & Arora, T. (2020). The relationship between sleep duration and mood in adolescents: A systematic review and meta-analysis. *Sleep Medicine Reviews, 52*(101311), 1-12. https://doi.org/10.1016/j.smrv.2020.101311

Short, M. A., Weber, N., Reynolds, C., Coussens, S., & Carskadon, M. A. (2018). Estimating adolescent sleep need using dose-response modeling. *Sleep, 41*(4), 1-14. https://doi.org/10.1093/sleep/zsy011

Sigurdson, J. F. (2019). Long-term effects of bullying in adolescence: Mental health and social adaptation in adulthood. [Doctoral dissertation, Norwegian University of Science and Technology]. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2596943

Silvers, J. A., McRae, K., Gabrieli, J. D., Gross, J. J., Remy, K. A., & Ochsner, K. N. (2012). Age-related differences in emotional reactivity, regulation, and rejection sensitivity in adolescence. *Emotion, 12*(6), 1235-1247. https://doi.org/10.1037/a0028297

Sisk, C. L., & Foster, D. L. (2004). The neural basis of puberty and adolescence. *Nature Neuroscience, 7*(10), 1040-1047. https://doi.org/10.1038/nn1326

Slavish, D. C., Sliwinski, M. J., Smyth, J. M., Almeida, D. M., Lipton, R. B., Katz, M. J., & Graham-Engeland, J. E. (2018). Neuroticism, rumination, negative affect, and sleep: Examining between-and within-person associations. *Personality and Individual Differences, 123*, 217-222. https://doi.org/10.1016/j.paid.2017.11.023

Smack, A. J., Kushner, S. C., & Tackett, J. L. (2015). Child personality moderates associations between parenting and relational and physical aggression. *Journal of Aggression, Maltreatment & Trauma*, *24*(7), 845-862. https://doi.org/10.1080/10926771.2015.1062450

Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression, 29*(3), 239-268. https://doi.org/10.1002/ab.10047

Somerville, L. H., Jones, R. M., & Casey, B. (2010). A time of change: behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. *Brain and Cognition, 72*(1), 124-133. https://doi.org/10.1016/j.bandc.2009.07.003

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science, 24*(8), 1554-1562. https://doi.org/10.1177/0956797613475633

Soto, C. J., & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality, 43*(1), 84-90. https://doi.org/10.1016/j.jrp.2008.10.002

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology, 94*(4), 718-737. https://doi.org/10.1037/0022-3514.94.4.718

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology, 100*(2), 330-348. https://doi.org/10.1037/a0021717

Soto, C. J., & Tackett, J. L. (2015). Personality traits in childhood and adolescence: Structure, development, and outcomes. *Current Directions in Psychological Science, 24*(5), 358-362. https://doi.org/10.1177/0963721415589345

Spear, L. P. (2009). Heightened stress responsivity and emotional reactivity during pubertal maturation: Implications for psychopathology. *Development and Psychopathology, 21*(1), 87-97. https://doi.org/10.1017/S0954579409000017

Spendelow, J. S., Simonds, L. M., & Avery, R. E. (2017). The relationship between co-rumination and internalizing problems: A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy, 24*(2), 512-527. https://doi.org/10.1002/cpp.2023

Spyropoulou, E., & Giovazolias, T. (2022). The scarring effects of depression symptoms on sadness rumination and peer victimization in Greek early adolescents: evidence for an integrated symptoms-driven model. *The Journal of Early Adolescence, 42*(7), 845-884. https://doi.org/10.1177/02724316221078827

Starr, L. R., & Davila, J. (2008). Excessive reassurance seeking, depression, and interpersonal rejection: a meta-analytic review. *Journal of Abnormal Psychology, 117*(4), 762-775. https://doi.org/10.1037/a0013866

Statista (2023). *Global smartphone sales to end users since 2007*. Retrieved 14th of August from https://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007/

Statistics Norway (2012). *Populations´ level of education*. Retrieved 16th of January from https://www.ssb.no/en/utdanning/statistikker/utniv/aar

Statistics Norway (SSB). (2021). *Kjønn, identitet og seksualitet i kvantitative spørreundersøkelser. Veileder for spørsmål om seksuell orientering, kjønnsidentitet, kjønnsuttrykk og kjønnskarakteristika, 2021/28*. Statistics Norway. Retrieved 11th of December from https://www.ssb.no/sosiale-forhold-og-kriminalitet/levekar/artikler/kjonn-identitet-og-seksualitet-i-kvantitative-sporreundersokelser/_/attachment/inline/86aeca7c-c6f5-4eee-a00a-8274c5261bb5:5c0f2c2442265393def722af81eca0530dba4539/NOT2021-28_web.pdf

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*(1), 78-106. https://doi.org/10.1016/j.dr.2007.08.002

Steinsbekk, S., Nesi, J., & Wichstrøm, L. (2023). Social media behaviors and symptoms of anxiety and depression. A four-wave cohort study from age 10–16 years. *Computers in Human Behavior, 147*(107859), 1-12. https://doi.org/10.1016/j.chb.2023.107859

Steinsbekk, S., & Wichstrøm, L. (2018). Cohort Profile: The Trondheim Early Secure Study (TESS)—a study of mental health, psychosocial development and health behaviour from preschool to adolescence. *International Journal of Epidemiology, 47*(5), 1401-1401i. https://doi.org/10.1093/ije/dyy190

Steinsbekk, S., Wichstrøm, L., Stenseng, F., Nesi, J., Hygen, B. W., & Skalická, V. (2021). The impact of social media use on appearance self-esteem from childhood to adolescence–A 3-wave community study. *Computers in Human Behavior, 114*(106528), 1-7. https://doi.org/10.1016/j.chb.2020.106528

Stockings, E., Degenhardt, L., Lee, Y. Y., Mihalopoulos, C., Liu, A., Hobbs, M., & Patton, G. (2015). Symptom screening scales for detecting major depressive disorder in children and adolescents: a systematic review and meta-analysis of reliability, validity and diagnostic utility. *Journal of Affective Disorders, 174*, 447-463. https://doi.org/10.1016/j.jad.2014.11.061

Stone, L. B., Hankin, B. L., Gibb, B. E., & Abela, J. R. (2011). Co-rumination predicts the onset of depressive disorders during adolescence. *Journal of Abnormal Psychology, 120*(3), 752-757. https://doi.org/10.1037/a0023384

Strøm, I. F., Aakvaag, H. F., Birkeland, M. S., Felix, E., & Thoresen, S. (2018). The mediating role of shame in the relationship between childhood bullying victimization and adult psychosocial adjustment. *European Journal of Psychotraumatology, 9*(1), 1-13. https://doi.org/10.1080/20008198.2017.1418570

Stumper, A., & Alloy, L. B. (2021). Associations between pubertal stage and depression: A systematic review of the literature. *Child Psychiatry & Human Development, 54*(2), 312-339. https://doi.org/10.1007/s10578-021-01244-0

Ståhl, S., & Dennhag, I. (2021). Online and offline sexual harassment associations of anxiety and depression in an adolescent sample. *Nordic Journal of Psychiatry, 75*(5), 330-335. https://doi.org/10.1080/08039488.2020.1856924

Sund, A. M., Larsson, B., & Wichstrøm, L. (2003). Psychosocial correlates of depressive symptoms among 12–14-year-old Norwegian adolescents. *Journal of Child Psychology and Psychiatry, 44*(4), 588-597. https://doi.org/10.1111/1469-7610.00147

Sveen, T. H., Berg-Nielsen, T. S., Lydersen, S., & Wichstrøm, L. (2013). Detecting psychiatric disorders in preschoolers: Screening with the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry, 52*(7), 728-736. https://doi.org/10.1016/j.jaac.2013.04.010

Sweeting, H., Young, R., West, P., & Der, G. (2006). Peer victimization and depression in early–mid adolescence: A longitudinal study. *British Journal of Educational Psychology, 76*(3), 577-594. https://doi.org/10.1348/000709905X49890

Swickert, R. J., Rosentreter, C. J., Hittner, J. B., & Mushrush, J. E. (2002). Extraversion, social support processes, and stress. *Personality and Individual Differences, 32*(5), 877-891. https://doi.org/10.1016/S0191-8869(01)00093-9

Tackett, J. L. (2006). Evaluating models of the personality–psychopathology relationship in children and adolescents. *Clinical Psychology Review, 26*(5), 584-599. https://doi.org/10.1016/j.cpr.2006.04.003

Tang, A.-M., Deng, X.-L., Du, X.-X., & Wang, M.-Z. (2018). Harsh parenting and adolescent depression: mediation by negative self-cognition and moderation by peer acceptance. *School Psychology International, 39*(1), 22-37. https://doi.org/10.1177/0143034317709066

The Norwegian Directorate for Children. (2017). *Barns Familier*. The Norwegian Directorate for Children. Retrieved 17th of October from https://www.bufdir.no/Statistikk_og_analyse/Oppvekst/Familie_omsorg_og_relasjoner/Barns_familier/

Thelle, D., & Laake, P. (2008). Epidemiologisk forskning: begreper og metoder. In P. Laake, B.R. Olsen & H. B. Benestad (Eds.), *Forskning i medisin og biofag* (pp. 282-320). Gyldendal Norsk Forlag.

Thompson, R. A., Flood, M. F., & Goodvin, R. (2006). Social support and developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Risk, disorder, and adaptation* (pp. 1–37). John Wiley & Sons, Inc.

Thorne, N., Witcomb, G. L., Nieder, T., Nixon, E., Yip, A., & Arcelus, J. (2019). A comparison of mental health symptomatology and levels of social support in young treatment seeking transgender individuals who identify as binary and non-binary. *International Journal of Transgenderism, 20*(2-3), 241-250. https://doi.org/10.1080/15532739.2018.1452660

Tian, L., Jiang, S., & Huebner, E. S. (2019). The big two personality traits and adolescents' complete mental health: The mediation role of perceived school stress. *School Psychology, 34*(1), 32-42. https://doi.org/10.1037/spq0000257

Tilton-Weaver, L., & Rose, A. (2023). The trade-offs of co-ruminating with friends: A profile analysis. *Journal of Research on Adolescence*. https://doi.org/10.1111/jora.12850

Twenge, J. M. (2020). Increases in depression, self-harm, and suicide among US adolescents after 2012 and links to technology use: possible mechanisms. *Psychiatric Research and Clinical Practice, 2*(1), 19-25. https://doi.org/10.1176/appi.prcp.20190015

Ullman, J. B. (2001). Structural Equation Modeling. In J. B. Ullman, B. G. Tabachnick, & L. S. Fidell (Eds). *Using multivariate statistics* (4th ed.) (pp. 653-771). Allyn & Bacon

Ullsperger, J. M., & Nikolas, M. A. (2017). A meta-analytic review of the association between pubertal timing and psychopathology in adolescence: Are there sex differences in risk? *Psychological Bulletin, 143*(9), 903-938. https://doi.org/10.1037/bul0000106

Vahedi, Z., & Zannella, L. (2021). The association between self-reported depressive symptoms and the use of social networking sites (SNS): *A meta-analysis. Current Psychology, 40,* 2174-2189. https://doi.org/10.1007/s12144-019-0150-6

Valkenburg, P. M., Beyens, I., Meier, A., & Abeele, M. M. V. (2022). Advancing our understanding of the associations between social media use and well-being. *Current Opinion on Psychology,* 47(101357), 1-7. https://doi.org/10.1016/j.copsyc.2022.101357

Van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related

validity study. *Journal of Research in Personality, 44*(3), 315-327. https://doi.org/10.1016/j.jrp.2010.03.003

Van Geel, M., Vedder, P., & Tanilon, J. (2014). Relationship between peer victimization, cyberbullying, and suicide in children and adolescents: a meta-analysis. *JAMA Pediatrics, 168*(5), 435-442. https://doi.org/10.1001/jamapediatrics.2013.4143

Vasey, M. W., Harbaugh, C. N., Lonigan, C. J., Phillips, B. M., Hankin, B. L., Willem, L., & Bijttebier, P. (2013). Dimensions of temperament and depressive symptoms: Replicating a three-way interaction. *Journal of Research in Personality, 47*(6), 908-921. https://doi.org/10.1016/j.jrp.2013.09.001

Vasileva, M., Graf, R. K., Reinelt, T., Petermann, U., & Petermann, F. (2021). Research review: A meta-analysis of the international prevalence and comorbidity of mental disorders in children between 1 and 7 years. *Journal of Child Psychology and Psychiatry, 62*(4), 372-381. https://doi.org/10.1111/jcpp.13261

Verhagen, M., Lodder, G. M., & Baumeister, R. F. (2018). Unmet belongingness needs but not high belongingness needs alone predict adverse well-being: A response surface modeling approach. *Journal of Personality, 86*(3), 498-507. https://doi.org/10.1111/jopy.12331

Verhulst, F. C., & Van der Ende, J. (2008). Using rating scales in a clinical context. In M. Rutter, D. V. M. Dorothy, M. A. Bishop, D. S. Pine, S. Scott, J. Stevenson, E. Taylor & A. Thapar (Eds.). *Rutter's Child and Adolescent Psychiatry* (5th ed.) (pp. 289-298). Wiley-Blackwell.

Vidal-Arenas, V., Bravo, A. J., Ortet-Walker, J., Ortet, G., Mezquita, L., Ibáñez, M. I., & Team, C.-c. A. S. (2022). Neuroticism, rumination, depression and suicidal ideation: A moderated serial mediation model across four countries. *International Journal of Clinical and Health Psychology, 22*(3), 1-10. https://doi.org/10.1016/j.ijchp.2022.100325

von Soest, T., Wichstrøm, L., & Kvalem, I. L. (2016). The development of global and domain-specific self-esteem from age 13 to 31. *Journal of Personality and Social Psychology, 110*(4), 592-608. https://doi.org/10.1037/pspp0000060

Waszczuk, M., Zavos, H., Gregory, A. M., & Eley, T. C. (2016). The stability and change of etiological influences on depression, anxiety symptoms and their co-occurrence across adolescence and young adulthood. *Psychological Medicine, 46*(1), 161-175. https://doi.org/10.1017/S0033291715001634

Weiss, B., & Garber, J. (2003). Developmental differences in the phenomenology of depression. *Development and Psychopathology, 15*(2), 403-430. https://doi.org/10.1017/S0954579403000221

Wendelborg, C. (2020). *Mobbing og arbeidsro i skolen: Analyse av Elevundersøkelsen skoleåret 2019/20.* NTNU Samfunnsforskning AS. https://samforsk.no/Publikasjoner/Mobbing-og-arbeidsro-i-skolen--Analyse-av-Elevundersøkelsen-skoleåret-20192020.aspx

Werner, A. M., Tibubos, A. N., Rohrmann, S., & Reiss, N. (2019). The clinical trait self-criticism and its relation to psychopathology: A systematic review–Update. *Journal of Affective Disorders, 246*, 530-547. https://doi.org/10.1016/j.jad.2018.12.069

Westenberg, P. M., Drewes, M. J., Goedhart, A. W., Siebelink, B. M., & Treffers, P. D. (2004). A developmental analysis of self-reported fears in late childhood through mid-adolescence: Social-evaluative fears on the rise? *Journal of Child Psychology and Psychiatry, 45*(3), 481-495. https://doi.org/10.1111/j.1469-7610.2004.00239.x

Wichstrøm, L. (1995). Harter's Self-Perception Profile for Adolescents: Reliability, validity, and evaluation of the question format. *Journal of Personality Assessment, 65*(1), 100-116. https://doi.org/10.1207/s15327752jpa6501_8

Wichstrøm, L. (1999). The emergence of gender difference in depressed mood during adolescence: the role of intensified gender socialization. *Developmental Psychology, 35*(1), 232-245. https://doi.org/10.1037/0012-1649.35.1.232

Wichstrøm, L., Belsky, J., & Steinsbekk, S. (2017). Homotypic and heterotypic continuity of symptoms of psychiatric disorders from age 4 to 10 years: a dynamic panel model. *Journal of Child Psychology and Psychiatry, 58*(11), 1239-1247. https://doi.org/10.1111/jcpp.12754

Wichstrøm, L., Berg-Nielsen, T. S., Angold, A., Egger, H. L., Solheim, E., & Sveen, T. H. (2012). Prevalence of psychiatric disorders in preschoolers. *Journal of Child Psychology and Psychiatry, 53*(6), 695-705. https://doi.org/10.1111/j.1469-7610.2011.02514.x

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), 10-18. https://doi.org/10.1111/j.1750-8606.2009.00110.x

Williams, A. L., Craske, M. G., Mineka, S., & Zinbarg, R. E. (2021). Neuroticism and the longitudinal trajectories of anxiety and depressive symptoms in older adolescents.

Journal of Abnormal Psychology, 130(2), 126-140. https://doi.org/10.1037/abn0000638

World Bank (2022). *Women, Business, and the Law 2022.* Washington, DC: World Bank. https://doi.org/10.1596/978-1-4648-1817-2

World Health Organization (2017). *Depression and other common mental disorders: global health estimates*. (Report no. WHO/MSD/MER/2017.2). World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf?sequence=1&isAllowed=y

World Health Organization (2022). *ICD-11: International classification of diseases (11th revision).* World Health Organization.

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M., Agerbo, Air, T. M., Andlauer, T. M. F., Becanu, S.-A., Bælvad-Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H., Bryois, J., Buttehschøn, H. N., Byjerg-Grauholm, J., . . . the Major Depressive Disorder Working Group of the Psyhciatric Genomic Consortium (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics, 50*(5), 668-681. https://doi.org/10.1038/s41588-018-0090-3

Xavier, A., Pinto Gouveia, J., & Cunha, M. (2016). Non-suicidal self-injury in adolescence: The role of shame, self-criticism and fear of self-compassion. *Child Youth Care Forum, 45,* 571-586. https://doi.org/10.1007/s10566-016-9346-1

Yang, H. J., Chiu, Y. J., Soong, W. T., & Chen, W. J. (2008). The roles of personality traits and negative life events on the episodes of depressive symptoms in nonreferred adolescents: a 1-year follow-up study. *Journal of Adolescent Health, 42*(4), 378-385. https://doi.org/10.1016/j.jadohealth.2007.09.017

Yoon, K. L., Maltby, J., & Joormann, J. (2013). A pathway from neuroticism to depression: Examining the role of emotion regulation. *Anxiety, Stress, & Coping, 26*(5), 558-572. https://doi.org/10.1080/10615806.2012.734810

Yoon, S., Kleinman, M., Mertz, J., & Brannick, M. (2019). Is social network site usage related to depression? A meta-analysis of Facebook–depression relations. *Journal of Affective Disorders, 248*, 65-72. https://doi.org/10.1016/j.jad.2019.01.026

Young, C. C., & Dietrich, M. S. (2015). Stressful life events, worry, and rumination predict depressive and anxiety symptoms in young adolescents. *Journal of Child and Adolescent Psychiatric Nursing, 28*(1), 35-42. https://doi.org/10.1111/jcap.12102

Zeigler-Hill, V., Holden, C. J., Enjaian, B., Southard, A. C., Besser, A., Li, H., & Zhang, Q. (2015). Self-esteem instability and personality: The connections between feelings of self-worth and the big five dimensions of personality. *Personality and Social Psychology Bulletin, 41*(2), 183-198. https://doi.org/10.1177/0146167214559719

Zhang, D., Huebner, E. S., & Tian, L. (2020). Longitudinal associations among neuroticism, depression, and cyberbullying in early adolescents. *Computers in Human Behavior, 112*, 1-9. https://doi.org/1016/j.chb.2020.106475

Zhu, X., Tian, L., & Huebner, E. S. (2019). Trajectories of suicidal ideation from middle childhood to early adolescence: risk and protective factors. *Journal of Youth and Adolescence, 48*, 1818-1834. https://doi.org/10.1007/s10964-019-01087-y

Zubin, J., & Spring, B. (1977). Vulnerability: a new view of schizophrenia. *Journal of Abnormal Psychology, 86*(2), 103-126. https://doi.org/10.1037/0021-843X.86.2.103

Study/Paper I

Morken, I. S., Viddal, K. R., Ranum, B., & Wichstrøm, L. (2020). Depression from preschool to adolescence—five faces of stability. *Journal of Child Psychology and Psychiatry*. *62*(8), 1000-1009. https://doi.org/10.1111/jcpp.13362

The Journal of Child Psychology and Psychiatry

Journal of Child Psychology and Psychiatry **:* (2020), pp **–** doi:10.1111/jcpp.13362

Depression from preschool to adolescence – five faces of stability

Ida Sund Morken,[1] Kristine Rensvik Viddal,[1] Bror Ranum,[1] and Lars Wichstrøm[1,2]

[1]Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway; [2]Department of Child and Adolescent Psychiatry, St. Olav's University Hospital, Trondheim, Norway

Background: The term 'stability' has different meanings, and its implications for the etiology, prevention, and treatment of depression vary accordingly. Here, we identify five types of stability in childhood depression, many undetermined due to a lack of research or inconsistent findings. **Methods:** Children and parents (n = 1,042) drawn from two birth cohorts in Trondheim, Norway, were followed biennially from ages 4–14 years. Symptoms of major depressive disorder (MDD) and dysthymia were assessed with the Preschool Age Psychiatric Assessment (only parents) and the Child and Adolescent Psychiatric Assessment (age 8 onwards). **Results:** (a) *Stability of form*: Most symptoms increased in frequency. The symptoms' importance (according to factor loadings) was stable across childhood but increased from ages 12–14, indicating that MDD became more coherent. (b) *Stability at the group level*: The number of symptoms of dysthymia increased slightly until age 12, and the number of symptoms of MDD and dysthymia increased sharply between ages 12–14. (c) *Stability relative to the group* (i.e., 'rank-order') was modest to moderate and increased from ages 12–14. (d) *Stability relative to oneself* (i.e., intraclass correlations) was stronger than stability relative to the group and increased from age 12–14. (e) *Stability of within-person changes*: At all ages, decreases or increases in the number of symptoms forecasted similar changes two years later, but more strongly so between ages 12–14. **Conclusions:** Across childhood, while most symptoms of MDD and dysthymia become more frequent, they are equally important. The transition to adolescence is a particularly vulnerable period: The depression construct becomes more coherent, stability increases, the level of depression increases, and such an increase predicts further escalation. Even so, intervention at any time during childhood may have lasting effects on reducing child and adolescent depression. **Keywords:** etiology; continuity; depression; developmental psychopathology; longitudinal.

Introduction

Depressive disorders are already present at preschool age and increase markedly during early adolescence (Merikangas et al., 2010; Wichstrøm et al., 2012). They then become one of the most prevalent mental health problems (World Health Organization, 2019). Given evidence showing stability in depression from childhood to adolescence, calls have been made for early prevention (e.g., Kovacs, Obrosky, & George, 2016), likely under the assumption that this stability indicates an etiological role of earlier depression in later depression. However, such an interpretation is not straightforward because available research leaves important questions unanswered: Do stability coefficients represent associations between the same phenomenon throughout development? Are reported coefficients (e.g., correlations), capturing a child's level of depression compared to other children, relevant when treating a specific child? What if stable underlying confounders produce the reported stability?

Prior developmental work has addressed distinct types of stability (e.g., Bornstein, Putnick, & Esposito, 2017; Roberts, Wood, & Caspi, 2008). However, research on developmental psychopathology has generally neglected the heterogeneous nature of stability, and studies on depression have been limited to stability of form, prevalence, and rank-order stability (e.g., Costello et al., 1996; Finsaas, Bufferd, Dougherty, Carlson, & Klein, 2018; Weiss & Garber, 2003). To achieve a more comprehensive understanding of depression as a developmental phenomenon and thereby inform prevention and treatment, we extend the focus to *within-person*-stability and change. We conceptualize and examine five types of stability: (a) form, (b) prevalence, (c) relative to the group, (d) relative to oneself, and (e) with respect to within-person changes, in symptoms of major depressive disorder (MDD) and dysthymia (subsumed under persistent depressive disorder), as defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013).

Stability of form

The symptoms that are most (a) *frequent* and (b) *important* to the depression construct (structural stability) define *stability of form*: whether depression presents similarly across developmental levels (Weiss & Garber, 2003). Although the frequency of individual symptoms may change with age, they could be equally important. Establishing structural stability is a prerequisite for interpreting other types

Conflict of interest statement: No conflicts declared.

Published by John Wiley & Sons Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main St, Malden, MA 02148, USA

of stability (e.g., stability coefficients). Notably, the specific criteria for childhood depression (irritable *or* depressed mood; 1-year duration of dysthymia as opposed to 2 years for adults) are identical *across* earlier childhood and adolescence, presupposing structural stability.

Meta-analytic evidence on (a) frequency of symptoms indicates that some symptoms (e.g., anhedonia, hopelessness) are more frequent in adolescents than in children (Weiss & Garber, 2003), but research after the mid-1990s is lacking. Evidence for (b) structural stability is inconsistent, as some studies report equal importance of symptoms (no change in factor loadings) (Weiss & Garber, 2003), whereas others chronicle altered importance of some symptoms (e.g., anhedonia increasingly important with age; Lahey et al., 2004; guilt more important in children, fatigue in adolescents; Weiss & Garber, 2003). Most nonclinical studies have used questionnaires that include items not corresponding to DSM-defined symptoms (e.g., 'I do everything wrong'), limiting conclusions that can be drawn about the structure of DSM-defined MDD and dysthymia.

Stability at the group level – prevalence

In short, questionnaire scores usually do not translate well into the number of DSM-defined symptoms or depressive disorders. Therefore, as regards prevalence, we focus on studies applying DSM criteria, typically assessed with clinical interviews. These findings reveal a low prevalence of MDD during the preschool years (e.g., 0.3%; Lavigne, LeBailly, Hopkins, Gouze, & Binns, 2009), and possibly higher prevalence in middle to late childhood (0.3%–1.6%; Ford, Goodman, & Meltzer, 2003; Merikangas et al., 2010) and early adolescence (1.9%–5.5%; Avenevoli, Swendsen, He, Burstein, & Merikangas, 2015; Merikangas et al., 2010). Dysthymia portrays a similar pattern, yet with lower prevalence and less of an increase; preschool 0.2%–0.3% (Lavigne et al., 2009; Wichstrøm et al., 2012); late childhood 0.8% (Merikangas et al., 2010), and early adolescence 0.2%–1.1% (Gau, Chong, Chen, & Cheng, 2005; Merikangas et al., 2010). Regarding any increases from late childhood to adolescence, two studies reported increase with respect to MDD (Ford et al., 2003; Merikangas et al., 2010), another did not (Costello et al., 1996), and two studies reported no change in dysthymia (Costello et al., 1996; Merikangas et al., 2010). The potential increase in MDD across earlier childhood has been examined by Ford et al. (2003), who concluded negatively. Dysthymia has not been examined.

Stability relative to the group

Traditionally, stability has been conceptualized as increased risk for future depression among those who have been depressed compared to those who have not, statistically examined by correlations, regressions, and odds ratios, which tap into an individual's risk of depression relative to the group. One study, which oversampled for psychiatric symptoms, found evidence for stability from preschool to adolescence when measuring depression *categorically* (Gaffrey, Tillman, Barch, & Luby, 2018), a finding also shown in clinical (e.g., Kovacs et al., 2016) but not community (Copeland et al., 2013) samples. However, given the low prevalence of depressive disorders in the general early childhood population, the latter null result may be due to limited power.

Two studies applying a *dimensional* approach – yielding more power – did indeed find evidence of stability in the community (Finsaas et al., 2018; Mason, Chmelka, Trudeau, & Spoth, 2017). Of note, Mason et al. (2017) studied stability from late childhood (age 11) and used a self-report questionnaire only partly covering DSM-defined symptoms of MDD and dysthymia. The Finsaas et al. (2018) study reported on combined MDD, dysthymia, and depressive disorder not otherwise specified. Given that the stability of different depressive disorders might differ (e.g., dysthymia enduring at least one year) and that one type of depression might increase the risk of another, the stability of combined depression may not transfer to each specific depression. Moreover, Finsaas et al. (2018) ended their study at age 12, thus not covering the expected steep increase in depression from around age 13 (e.g., Ford et al., 2003). In sum, the magnitude of 'rank-order' stability of DSM-5-defined MDD and dysthymia in the population during early childhood is uncertain, and the stability into adolescence remains unknown.

Stability relative to oneself

For treatment and indicated prevention, knowing to what extent a child maintains his or her depression score relative to others may be of limited value. This is especially relevant at ages when the population average changes, as in early adolescence. A more clinically relevant question would be 'How likely is this child to be *equally* depressed in the future?'. This within-person-type question cannot be accurately answered by approaches mixing between-person and within-person information (e.g., correlations). No previous study has conceptualized and investigated the within-person stability of childhood depression.

Stability of within-person change

To what extent do stability findings implicate a causal relation between early and later depression? Indeed, the portrayed stability might be produced by third variables. Within-person approaches, which examine whether change predicts further change using the person as her or his own control, account for one class of potential confounders; those that do

not change over time (e.g., genetics, Wray et al., 2018). However, only one study, applying a dynamic panel model, has utilized children as their own controls (Wichstrøm, Belsky, & Steinsbekk, 2017). The age-range was limited to 10 years and younger, while we include adolescents.

Current study

In sum, the term 'stability' may be understood in a variety of ways – with differing implications for therapy and prevention. Here, childhood depression will be investigated with respect to (a) stability of form, (b) stability at the group level (prevalence), (c) stability of the number of symptoms relative to the group, (d) relative to oneself, and (e) stability of within-person changes in the number of symptoms. By extending the focus to *within-person* stability and change, we provide a novel approach to studying developmental psychopathology – beyond prevalence and rank-order stability.

Methods

Participants and procedure

The Trondheim Early Secure Study (TESS; Wichstrøm et al., 2012) comprises members of the 2003 and 2004 birth cohorts in Trondheim, Norway (N = 3,456). A letter of invitation along with the Strengths and Difficulties Questionnaire (SDQ) version 4–16 (Goodman, Ford, Simmons, Gatward, & Meltzer, 2000) was sent to the children's homes. Almost all children met at the age-4 routine health check-up (n = 3,358). Parents were informed about the TESS by a health nurse, using procedures approved by the Regional Committee for Medical and Health Research Ethics, Mid-Norway, and written consent was obtained. To increase statistical power, children were divided into four strata based on their SDQ score (0–4, 5–8, 9–11, 12–40), and the probability of being selected increased with increased SDQ scores (37%, 48%, 70%, and 89% from the respective strata). This oversampling of mental health problems was accounted for in the analyses. The drop-out rate after consent at the well-child clinic did not differ across SDQ score or gender. Of the 1,250 children randomly selected for the Study, 997 were successfully enrolled at Time 1 (M_{age} = 4.7 years, SD = .30; 49.1% boys). Retesting occurred biennially at 6, 8, 10, 12, and 14 years (M_{age} = 6.7, SD = 0.17; M_{age} = 8.8, SD = 0.24; M_{age} = 10.5, SD = 0.15; M_{age} = 12.5, SD = 0.15; M_{age} = 14.4, SD = 0.16, respectively, Figure S1). Participants with information from at least one wave of data collection comprised the analytical sample (n = 1,042).

More symptoms of depression at some waves predicted reduced retention at some of the later waves: age-6 MDD symptoms to age-8 retention (OR = 0.82, 95% CI:0.67, 1.00); age-12 MDD symptoms to age-14 retention (OR = 0.72, CI:0.59, 0.87); age-4 dysthymia symptoms to age-12 retention (OR = 0.84, CI:0.70, 1.00), age-6 dysthymia symptoms to age-10 (OR = 0.78, CI:0.63, 0.97), and age-12 retention (OR = 0.80, CI:0.65, 0.98); and age-12 dysthymia symptoms to age-14 retention (OR = 0.74, CI:0.61, 0.89).

Measures

The Preschool Age Psychiatric Assessment (PAPA) and the Child and Adolescent Psychiatric Assessment (CAPA). Symptoms of MDD and dysthymia were measured using the semi-structured psychiatric interviews PAPA (Egger & Angold, 2006; ages 4 and 6; parents) and CAPA (Angold & Costello, 2000; age 8 and onwards). In the PAPA, persistent preoccupation with death or self-harm in play were, as proposed by Luby et al. (2002), added as age-appropriate criteria for suicidality. Children and parents were interviewed separately with the CAPA, and a symptom was considered present if reported by either respondent. Interviewers (n = 7) had at least a bachelor's degree in relevant fields, substantial experience with children and families and were trained by the team who developed the PAPA/CAPA. Inter-rater reliabilities among blinded coders of 9% of videotapes of PAPA interviews and 15% of audiotapes of CAPA interviews were .91 and .87, respectively, for MDD symptoms, and .89 and .85, respectively, for dysthymia symptoms.

Statistical analyses

Analyses were performed in Mplus 8.1 using a robust maximum likelihood estimator and probability weights to correct for the oversampling of children with mental health problems. Missing data were handled using a full information maximum likelihood (FIML) procedure.

The stability of form was examined in two ways: (a) the frequency of symptoms with percentages by age group, and linear and quadratic latent growth curves, and (b) structural stability (metric equivalence) by confirmatory factor analysis (CFA), comparing the model fit when all factor loadings were freely estimated to when they were fixed to be equivalent of two adjoining ages (Widaman, Ferrer, & Conger, 2010). Due to categorical variables, akaike information criterion (AIC) $\geq$ 2 was used as the criterion (Burnham & Anderson, 2004).

Group-level stability was examined descriptively by percentages receiving the diagnosis of MDD and dysthymia and by mean levels of symptoms. Changes in the number of symptoms were examined by latent growth curves, considering symptom counts to follow a negative binominal distribution.

Stability relative to the group was investigated with Pearson's r. To test whether stability increased or decreased with age, we compared models where correlations between adjacent ages were fixed to be equivalent as opposed to freely estimated, using the Satorra–Bentler scaled chi-square test.

Stability relative to oneself (individuals keeping their depression level over time) was investigated with the intraclass correlation coefficient (ICC), capturing the proportion of the total variance attributed to between-person variance.

The stability of within-person changes was examined with random intercept cross-lagged panel model (RI-CLPM). In this model, within-person variance is separated from between-person variance (Hamaker, Kuiper, & Grasman, 2015), and thus, we estimate whether deviations from one's mean level of depression predict later deviations, while ruling out time-invariant confounders.

Results

Stability of form

Across ages 4-14, the most frequent symptoms of MDD were irritability, change in weight/appetite, sleep disturbances, and diminished concentration (Table 1). The most frequent symptoms of dysthymia were changes in weight/appetite, sleep disturbances, and diminished concentration (also symptoms of MDD) and low self-esteem. Most symptoms increased in a linear and/or a quadratic fashion. Although low self-esteem increased linearly, it

Table 1 Frequency and mean level change of individual symptoms of MDD and Dysthymia, age 4–14

	Frequency % [95% CI]						Mean level change	
	Age 4	Age 6	Age 8	Age 10	Age 12	Age 14	Linear slope [95% CI], *p*-value	Quadratic slope [95% CI], *p*-value
MDD symptoms								
1a. Depressive Mood*	0.7 [0.2, 1.1]	2.1 [1.0, 3.2]	1.8 [0.7, 2.8]	2.4 [1.1, 3.7]	3.4 [1.9, 4.9]	7.8 [5.5, 10.1]	0.42 [0.27, 0.56], $p < .001$	0.05 [0.01, 0.10], $p = .019$
1b. Irritable Mood*	13.6 [11.4, 15.7]	16.8 [14.0, 19.5]	9.3 [7.0, 11.6]	15.0 [12.1, 17.9]	17.8 [14.7, 20.9]	17.7 [14.5, 20.9]	0.45 [0.12, 0.78], $p = .008$	0.14 [0.03, 0.25], $p = .013$
2. Adhedonia	0.5 [0.1, 0.9]	1.2 [0.2, 2.2]	0.8 [0.1, 1.5]	1.4 [0.4, 2.4]	2.1 [2.2, 5.3]	4.3 [2.6, 6.0]	0.22 [0.11, 0.33], $p < .001$	0.04 [0.01, 0.08], $p = .023$
3. Weight/ Appetite*	8.7 [6.8, 10.6]	5.8 [4.1, 7.6]	4.9 [3.3, 6.6]	6.1 [4.3, 8.0]	9.9 [7.6, 12.3]	12.5 [9.7, 15.4]	0.37 [0.12, 0.63], $p = .004$	0.21 [0.13, 0.30], $p < .001$
4. Sleep disturbances*	12.3 [10.1, 14.5]	10.0 [7.8, 12.2]	10.8 [8.4, 13.2]	9.3 [7.1, 11.6]	6.7 [4.7, 8.7]	6.7 [4.6, 8.8]	−0.57 [−0.83, −0.32], $p < .001$	−0.01 [−0.08, 0.06], $p = .793$
5. Agitation/ Retardation	0.5 [0.1, 1.0]	2.0 [0.9, 3.0]	0.0 [0.0, 0.1]	–[a]	0.5 [−0.2, 1.1]	1.6 [0.6, 2.7]	0.01 [−0.02, 0.04], $p = .337$	0.04 [0.02, 0.06], $p = .001$
6. Fatigue/Energy loss*	1.0 [0.3, 1.6]	3.0 [1.7, 4.2]	1.6 [0.6, 2.5]	5.7 [3.9, 7.5]	7.6 [5.5, 9.8]	14.4 [11.4, 17.4]	.083 [0.63, 1.03], $p < .001$	0.16 [0.11, 0.22], $p < .001$
7. Guilt/ Worthlessness	2.0 [1.2, 2.7]	7.5 [5.6, 9.5]	8.7 [6.6, 10.9]	14.7 [12.0, 17.5]	11.1 [8.6, 13.6]	14.2 [11.3, 17.2]	1.34 [1.05, 1.63], $p < .001$	−0.14 [−0.22, 0.06], $p = .001$
8. Concentration*	7.4 [5.8, 9.0]	11.6 [9.3, 13.9]	16.7 [13.7, 19.6]	13.5 [10.9, 16.2]	17.7 [14.6, 20.9]	19.0 [15.7, 22.3]	1.19 [0.85, 1.52], $p < .001$	−0.08 [−.018, 0.02], $p = .109$
9. Suicidality	5.4 [3.9, 7.0]	4.7 [3.1, 6.3]	1.1 [0.2, 1.9]	1.4 [0.4, 2.4]	1.4 [0.4, 2.4]	1.6 [0.6, 2.7]	−.027 [−0.42, −0.13], $p < .001$	0.10 [0.05, 0.15], $p = .252$
Dysthymia symptoms								
5. Low self-esteem	0.4 [0.1, 0.8]	4.4 [2.9, 5.9]	7.6 [5.6, 9.6]	13.2 [10.6, 15.9]	8.9 [6.6, 11.2]	11.7 [9.0, 14.4]	1.34 [1.10, 1.60], $p < .001$	−0.14 [−0.21, −0.07], $p < .001$
7. Hopelessness	0.1 [−.01, 0.2]	0.3 [−0.1, 0.6]	0.6 [0.1, 1.1]	1.7 [0.6, 2.7]	2.5 [1.2, 3.7]	4.6 [2.8, 6.5]	0.27 [0.17, 0.37], $p < .001$	0.05 [0.02, 0.07], $p = .001$

[a]The values lack due to low frequency.
*Overlapping symptoms of MDD and dysthymia.

evinced a quadratic decrease, possibly given a peak at age 10. Furthermore, sleep disturbances and suicidality decreased linearly. Notably, death themes in play – specific to the suicidality symptom at ages 4 and 6 – were infrequent (1.7% and 0.5%, respectively).

The results from CFAs indicated structural stability of MDD and dysthymia at most ages (Table 2). For dysthymia, the equality of factor loadings evinced *better* fit than freely estimated loadings ($\Delta AIC_{4-6} = -20.83$, $\Delta AIC_{6-8} = -1.87$, $\Delta AIC_{8-10} = -10.47$, $\Delta AIC_{10-12} = -1.96$, $\Delta AIC_{12-14} = -2.11$). For MDD, the equality of factor loadings up to age 12 also demonstrated superior fit ($\Delta AIC_{4-6} = -482.93$, $\Delta AIC_{6-8} = -10.27$, $\Delta AIC_{8-10} = -9.88$, $\Delta AIC_{10-12} = -5.74$), whereas the reverse was seen from ages 12–14, as indicated by a better model fit when symptoms were freely estimated ($\Delta AIC_{12-14} = 8.93$) reflecting higher loadings at age 14.

Stability at the group level – prevalence

As regards number of symptoms, beyond linear increases in MDD and dysthymia from ages 4–14, the number of MDD symptoms increased curvilinearly (Table 3). From ages 4–12, there was neither a linear ($slope_l = .00$, $p = .980$) nor a quadratic ($slope_q = .00$, $p = .380$) increase in MDD symptoms, indicating that the curvilinearity was due to the increase from ages 12–14. In contrast, the number of symptoms of dysthymia increased from ages 4–12 ($slope_l = .03$, $p < .001$).

At a diagnostic level, the prevalence of MDD appeared lower in preschool than adolescence (Table 3), but a growth curve could not be estimated due to the very low prevalence at ages 8 and 10. The prevalence of dysthymia diagnosis was somewhat higher, evincing a curvilinear increase from ages 4–14. Childhood dysthymia diagnosis (i.e., ages 4–12) increased linearly ($slope_l = .31$, $p = .002$), but not quadratically ($slope_q = .01$, $p = .600$), indicating that the curvilinearity for the whole period was due to the sharp increase from ages 12–14.

Stability relative to others

Pearson's correlation analyses revealed modest to moderate stability in the number of symptoms of MDD and dysthymia relative to the other participants' symptoms, typically $r = .30–.35$ (Table S1, upper triangle).

The stability from age 4 to later ages, however, was low and tapered off from age 10 onwards. From ages 12–14, the stability was stronger than in the 10–12 age span, both in MDD ($\Delta\chi^2 = 32.54$, $p < .001$) and dysthymia ($\Delta\chi^2 = 20.17$, $p < .001$). In the younger age-groups, no change in stability was observed for MDD (8–10 versus 10–12 $\Delta\chi^2 = .00$, $p = 1.00$; 6–8 versus 8–10 $\Delta\chi^2 = .01$, $p = .943$; 4–6 versus 6–8 $\Delta\chi^2 = .44$, $p = .507$) or dysthymia (8–10 versus 10–12 $\Delta\chi^2 = .00$, $p = 1.00$; 6–8 versus 8–10 $\Delta\chi^2 = .56$, $p = .453$; 4–6 versus 6–8 $\Delta\chi^2 = 1.57$, $p = .194$).

Stability relative to oneself

ICCs were of similar magnitude as Pearson's correlations (Table S1, lower triangle). Note that the ICC is a direct expression of explained variance, whereas a similar expression for the Pearson's correlation is r^2. Across ages, the average r^2 was .09 for MDD and .08 for dysthymia, while the average ICC was .22 for MDD and .21 for dysthymia. Thus, the stability relative to oneself was more than twice as strong as the relative to the group. Tables S2–S6 show that in early and middle childhood (i.e., ages 4–10), most children with depressive symptoms had fewer symptoms two years later, whereas approximately ¼ of those without symptoms had acquired some symptoms. Those with symptoms at age 12 were, however, more often inclined to remain at the same level or to have an increase in their number of symptoms at age 14.

Stability of within-person change

An increased or reduced number of symptoms, of both MDD and dysthymia – and at all ages, forecasted a corresponding increase or decrease two years later (Table 4). In addition, for both MDD and dysthymia, within-person changes at age 10 predicted similar changes at age 14 over and above the changes at age 12. Furthermore, within-person changes in symptoms of dysthymia at age 8 predicted similar changes at age 12 and over and above the changes at age 10. The strength of the two-year prediction was stronger for the 12–14-year span than for the 10-12 span for both MDD ($\Delta\chi^2 = 32.54$, $p < .001$) and dysthymia ($\Delta\chi^2 = 20.17$, $p < .001$). At younger ages, strengths of predictions were of similar magnitudes: MDD (4–6 versus 6–8 $\Delta\chi^2 = 0.44$, $p = .51$; 6–8 versus 8–10 $\Delta\chi^2 = 0.01$, $p = .94$; 8–10 versus 10–12 $\Delta\chi^2 = 0.00$, $p = 1.00$); dysthymia (4–6 versus 6–8 $\Delta\chi^2 = 1.57$, $p = .194$; 6–8 versus 8–10 $\Delta\chi^2 = 0.56$, $p = .453$; 8–10 versus 10–12 $\Delta\chi^2 = 0.00$, $p = 1.00$).

Discussion

We investigated five types of stability in symptoms of MDD and dysthymia by following a representative community sample from preschool to adolescence. Across childhood, while most symptoms increased, their importance did not change. When entering adolescence, the depression construct became more coherent, there was an increase in prevalence, stability relative to oneself and others, and stability of within-person changes.

Table 2 Factor loadings for MDD and Dysthymia, age 4–14

	Standardized factor loadings					
	Age 4	Age 6	Age 8	Age 10	Age 12	Age 14
MDD symptoms						
1a. Depressive Mood	.05	.61***	.71***	.62***	.71***	.96***
1b. Irritable Mood	.97***	.96***	.62***	.62***	.65***	.57***
2. Adhedonia	.43	.24	.74***	.62***	.65***	.69***
3. Weight/Appetite	.26**	.08	.46***	.61***	.58***	.50***
4. Sleep disturbance	.33**	.45***	.39***	.46***	.58***	.74***
5. Agitation/Retardation	.49*	.39*	.26	_[a]	_[a]	.79***
6. Fatigue/Energy loss	.47**	.45***	.58****	.62***	.75***	.80***
7. Guilt/Worthlessness	.38*	.41***	.78***	.69***	.62***	.80***
8. Concentration	.39	.40***	.53***	.46***	.39***	.68***
9. Suicidality	.08	.37***	.58**	.79***	.79***	.69***
Dysthymia symptoms						
1. Depressed/Irritable Mood	.31*	.32**	.72***	.64***	.83***	.82***
2. Weight/Appetite	.22	.28*	.47***	.57***	.43***	.52***
3. Sleep disturbance	.49***	.49***	.45***	.36***	.59***	.81***
4. Fatigue/energy loss	.64***	.48*	.60***	.57***	.66***	.79***
5. Low self-esteem	.64***	.51***	.76***	.77***	.87***	.80***
6. Concentration	.63***	.64***	.57***	.45***	.52***	.69***
7. Hopelessness	.98***	.64	.60*	.83***	.88***	.80***

[a]The values lack due to low prevalence.
*Indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$.

Stability of form

Frequency of individual symptoms. Depressed mood and anhedonia were fairly low frequent, whereas irritability was much more frequent. Most symptoms increased linearly or even quadratically, whereas some decreased. Somewhat surprisingly, suicidality was highest among 4- and 6-year-olds. Conceivably, this could have been caused by frequent death themes in play; however, this was very low frequent. A more probable explanation is preschoolers asking parents many questions, including questions about death (e.g., Chouinard, Harris, & Maratsos, 2007). The relatively low and stable frequency from age 8 to 14 contradicts commonly reported findings of increased suicidality in adolescence (Kennebeck & Bonin, 2017).

Structural stability. Previous findings on structural stability from childhood to adolescence are inconsistent (Lahey et al., 2004; Weiss & Garber, 2003). Here, we found that although the frequency of symptoms changed, their importance to MDD and dysthymia did not change across childhood, indicating an isomorphic depression construct over time. Thus, the reported stability coefficients represent associations between the same phenomenon, as presupposed in the DSM-5. However, we did not investigate potential symptoms not listed in the DSM-5 and discrimination from symptoms of other DSM-5 diagnoses, and hence, the present findings do not inform on the developmental *appropriateness* of the DSM-5 criteria.

Although factor loadings did not change during childhood, increased loadings from ages 12-14 indicate that MDD becomes an even more coherent construct in adolescence than in childhood. Thus, before adolescence – and when taking the increased prevalence into account, if one symptom of MDD emerges, it is typically limited to that symptom, or a few others. By contrast, when an adolescent is evincing the same symptom, other symptoms tend to be present as well, likely increasing the odds of a full-blown MDD diagnosis. Thus, when an adolescent is demonstrating a symptom of depression, further clinical assessment is especially warranted.

Stability at the group level – prevalence

Major depressive disorder, both the number of symptoms and diagnosis, was low-prevalent from preschool to preadolescence, which is in line with Ford et al. (2003) finding diagnosed MDD low-prevalent in these age-groups. We also identified a sharp increase in the number of MDD symptoms from ages 12–14, in line with previous findings showing an increase in MDD diagnosis in early adolescence (Ford et al., 2003; Merikangas et al., 2010). Notably, Costello et al. (1996) did *not* find an increase in diagnosed MDD from ages 11–13. Hence, it is possible that the increase first appears between ages 13–14.

The present study is the first to investigate the stability of dysthymia in early childhood, finding an increase in prevalence, both number of symptoms and diagnosis, from ages 4–12. At a diagnostic level, we found an even accelerated increase from ages 12–14, contrasting studies showing no change in prevalence from middle childhood to adolescence (Costello et al., 1996; Merikangas et al., 2010). As diagnosed dysthymia in the present study was 2–5 times higher in early adolescence than in reports from the United

Table 3 Prevalence of diagnosis, mean level of symptoms, and mean level change of MDD and Dysthymia, age 4–14

	Prevalence of diagnosis % [95% CI]						Mean level change	
	Age 4	Age 6	Age 8	Age 10	Age 12	Age 14	Linear slope [95% CI], *p*-value	Quadratic slope [95% CI], *p*-value
MDD	0.11 [−0.04, 0.26]	0.33 [−0.20, 0.85]	_[a]	_[a]	0.71 [−0.17, 1.60]	2.68 [1.02, 4.33]	_[a]	_[a]
Dysthymia	0.39 [−0.10, 0.88]	0.40 [−0.14, 0.93]	0.83 [0.02, 1.63]	0.67 [−0.11, 1.44]	2.02 [0.58, 3.47]	5.08 [2.81, 7.35]	0.3 [0.01, 0.59], $p = .044$	0.22 [0.01, 0.43], $p = .041$
Mean level of symptoms (SD)								
MDD	0.75 (0.04)	0.83 (0.05)	0.69 (0.05)	0.79 (0.05)	0.91 (0.07)	1.11 (0.09)	0.03 [0.01, 0.04], $p < .001$	0.01 [0.00, 0.01], $p < .001$
Dysthymia	0.44 (0.03)	0.50 (0.04)	0.59 (0.04)	0.62 (0.04)	0.66 (0.05)	.81 (0.07)	0.03 [0.02, 0.05], $p < .001$	0.00 [0.00, 0.00], $p = .530$

[a]The values lack due to low prevalence.

States (Costello et al., 1996; Merikangas et al., 2010), we possibly had more power to detect increases.

Stability relative to others and to oneself

Pearson's correlations portrayed moderate to low stability of depressive symptoms from age 6 to all later time points until age 14. However, from age 4, there was stability only until age 6 and 8. Overall, this modest 'rank-order' stability of depressive symptoms across childhood is buttressed by reports from previous population studies (Finsaas et al., 2018; Mason et al., 2017).

The stability relative to oneself was moderate to moderately high and higher than that relative to others. Thus, if one is interested in the proclivity of children to evince the same symptom load at later time points and lean on the commonly reported correlation coefficients to answer this question, stability will likely be underestimated. At closer inspection, instability in early and middle childhood was due to some children evincing reductions and others evincing increases in the number of symptoms. When entering adolescence, instability due to reduced number of symptoms diminished. In other words, compared to earlier periods, depressed early adolescents less often outgrow their depression in mid-adolescence. Thus, interventions may be particularly important when entering adolescence.

Stability of within-person change

We found, for the first time that, throughout childhood, changes in depression predicting later changes in the same direction, even more strongly from age 12–14. We cannot rule out that time-*varying* factors (e.g., negative life events, hormonal changes) produced these results. However, they are consistent with the view that having a depressive episode – or even subclinical depression – could be part of the etiology of persistent or recurring depression (Rohde, Lewinsohn, & Seeley, 1994). If this holds true, the treatment of childhood depression resulting in symptom reduction should have lasting effects – even when not altering the original etiological factors. A range of 'scar'-like mechanisms are possible, both intrapersonal, such as structural or functional brain alterations (e.g., reduced hippocampal volume; Mikolas et al., 2018) and interpersonal, such as social withdrawal discouraging age-mates from interacting with the child (Thompson, Flood, & Goodvin, 2006). Identifying and targeting such maintaining mechanisms could help prevent persistent or recurrent depression.

Limitations

Although the present study has a range of strengths, including investigating five types of stability,

Table 4 Random intercept analysis of symptoms of MDD and Dysthymia, age 4–14

	Standardized slope coefficients (*p*-value)				
MDD					
Two-year span	Age 4–6: .06 ($p = .486$)	Age 6–8: .14 ($p = .049$)	Age 8–10: .13 ($p = .037$)	Age 10–12: .25 ($p < .001$)	Age 12–14: .46 ($p < .001$)
Four-year span	Age 4–8: .10 ($p = .188$)	Age 6–10: .08 ($p = .104$)	Age 8–12: .13 ($p = .054$)	Age 10–14: .21 ($p = .002$)	
Dysthymia					
Two-year span	Age 4–6: .07 ($p = .493$)	Age 6–8: .19 ($p = .010$)	Age 8–10: .21 ($p = .001$)	Age 10–12: .28 ($p < .001$)	Age 12–14: .47 ($p < .001$)
Four-year span	Age 4–8:.10 ($p = .229$)	Age 6–10: .05 ($p = .300$)	Age 8–12: .15 ($p = .018$)	Age 10–14: .18 ($p = .002$)	

following a community sample over many waves from preschool to adolescence, and studying both DSM-5-defined MDD and dysthymia – with clinical interviews – we acknowledge several limitations. First, we mainly studied symptom counts. There is no compelling evidence pointing to depression being categorical in nature (Haslam, Holland, & Kuppens, 2012), yet we cannot be sure that our findings apply to diagnosed disorders. Second, although we adjusted for all time-invariant confounders, time-varying factors may have influenced the reported stability coefficients. Third, some symptoms of depression were infrequent, resulting in wide confidence intervals of factor loadings and frequency rates. Fourth, depressed participants more often dropped out of the study, likely underestimating the increase in prevalence. However, the use of an FIML approach to missingness likely mitigated some of this selective attrition. Fifth, because a 3-month primary period was applied, depressive symptoms occurring between our 2-year intervals of observation might have been missed. Sixth, given our Norwegian sample, generalization to other cultures/countries should be performed with caution.

Conclusions

This investigation highlighted five faces of the stability of depression. Our results on the stability of form indicate that while most symptoms of MDD and dysthymia increase – and some decrease – in frequency during childhood and early adolescence, their importance to the depression constructs do not change. Regarding group-level stability, whereas dysthymia increases slightly from ages 4–12 – symptoms of both disorders increase sharply when entering adolescence. The stability of one's absolute level of depressive symptoms (stability relative to oneself) is moderate, but stronger than the modest stability portrayed by 'rank-order' correlations (stability relative to the group). When entering adolescence, the depression construct becomes more coherent, prevalence and rank-order stability increases, and symptoms are less likely to remit than before. These findings suggest early adolescence as a particularly vulnerable period. Importantly though, changes in depression predicting later changes in depression (ruling out time-invariant confounding) at all ages – but especially in adolescence, holds promise for early intervention to prevent recurrent, persistent, or increased depression.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Table S1. Stability relative to others (Pearson's correlation coefficient) stability relative to others (Intra-class correlation coefficient) between number of symptoms of MDD and Dysthymia, age 4–14.
Table S2. Percentages with symptoms at age 6 sorted by symptoms at age 4.
Table S3. Percentages with symptoms at age 8 sorted by symptoms at age 6.
Table S4. Percentages with symptoms at age 10 sorted by symptoms at age 8.
Table S5. Percentages with symptoms at age 12 sorted by symptoms at age 10.
Table S6. Percentages with symptoms at age 14 sorted by symptoms at age 12.
Figure S1. Flow chart of recruitment and follow-up.

Acknowledgements

The authors have declared that they have no competing or potential conflicts of interest.

Correspondence

Ida Sund Morken, Department of Psychology, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway; Email: ida.sund.morken@gmail.-com

Key points

- Prior research has reported modest to moderate rank-order stability of depression in childhood and adolescence. Other types of stability have seldom or – for some types – not been examined.
- The stability of childhood depression relative to oneself is stronger than relative to others (rank order). Hence, if one is interested in the proclivity of children to evince the same symptom load at later time points, this type stability will be underestimated by relying on rank-order measures; stability relative to oneself should be used instead.
- Improvement or deterioration in depression predicts corresponding changes in later depression when all time-invariant confounding is adjusted for. Thus, early intervention may prevent recurrent, persistent, or increased depression.

References

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). *American Psychiatric Pub*. ISO 690.

Angold, A., & Costello, E.J. (2000). The child and adolescent psychiatric assessment (CAPA). *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 39–48.

Avenevoli, S., Swendsen, J., He, J.-P., Burstein, M., & Merikangas, K.R. (2015). Major depression in the national comorbidity survey–adolescent supplement: Prevalence, correlates, and treatment. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*, 37–44. e32.

Bornstein, M.H., Putnick, D.L., & Esposito, G. (2017). Continuity and stability in development. *Child Development Perspectives*, *11*, 113–119.

Burnham, K.P., & Anderson, D.R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*, 261–304.

Chouinard, M.M., Harris, P.L., & Maratsos, M.P. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, *72*, i–i129.

Copeland, W.E., Adair, C.E., Smetanin, P., Stiff, D., Briante, C., Colman, I., … & Angold, A. (2013). Diagnostic transitions from childhood to adolescence to early adulthood. *Journal of Child Psychology and Psychiatry*, *54*, 791–799.

Costello, E.J., Angold, A., Burns, B.J., Stangl, D.K., Tweed, D.L., Erkanli, A., & Worthman, C.M. (1996). The Great Smoky Mountains Study of Youth: Goals, design, methods, and the prevalence of DSM-III-R disorders. *Archives of General Psychiatry*, *53*, 1129–1136.

Egger, H.L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry*, *47*, 313–337.

Finsaas, M.C., Bufferd, S.J., Dougherty, L.R., Carlson, G.A., & Klein, D.N. (2018). Preschool psychiatric disorders: homotypic and heterotypic continuity through middle childhood and early adolescence. *Psychological Medicine*, *48*, 2159–2168.

Ford, T., Goodman, R., & Meltzer, H. (2003). The British child and adolescent mental health survey 1999: The prevalence of DSM-IV disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*, 1203–1211.

Gaffrey, M.S., Tillman, R., Barch, D.M., & Luby, J.L. (2018). Continuity and stability of preschool depression from childhood through adolescence and following the onset of puberty. *Comprehensive Psychiatry*, *86*, 39–46.

Gau, S.S., Chong, M.-Y., Chen, T.H., & Cheng, A.T. (2005). A 3-year panel study of mental disorders among adolescents in Taiwan. *American Journal of Psychiatry*, *162*, 1344–1350.

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry*, *177*, 534–539.

Hamaker, E.L., Kuiper, R.M., & Grasman, R.P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*, 102.

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychological Medicine*, *42*, 903–920.

Kennebeck, S., & Bonin, L. (2017). Suicidal behavior in children and adolescents: Epidemiology and risk factors. *UptoDate* [online Database]. Last Updated, 21.

Kovacs, M., Obrosky, S., & George, C. (2016). The course of major depressive disorder from childhood to young adulthood: Recovery and recurrence in a longitudinal observational study. *Journal of Affective Disorders*, *203*, 374–381.

Lahey, B.B., Applegate, B., Waldman, I.D., Loft, J.D., Hankin, B.L., & Rick, J. (2004). The structure of child and adolescent psychopathology: generating new hypotheses. *Journal of Abnormal Psychology*, *113*, 358.

Lavigne, J.V., LeBailly, S.A., Hopkins, J., Gouze, K.R., & Binns, H.J. (2009). The prevalence of ADHD, ODD, depression, and anxiety in a community sample of 4-year-olds. *Journal of Clinical Child & Adolescent Psychology*, *38*, 315–328.

Luby, J.L., Heffelfinger, A.K., Mrakotsky, C., Hessler, M.J., Brown, K.M., & Hildebrand, T. (2002). Preschool major depressive disorder: preliminary validation for developmentally modified DSM-IV criteria. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*, 928–937.

Mason, W.A., Chmelka, M.B., Trudeau, L., & Spoth, R.L. (2017). Gender moderation of the intergenerational transmission and stability of depressive symptoms from early adolescence to early adulthood. *Journal of Youth and Adolescence*, *46*, 248–260.

Merikangas, K.R., He, J.-P., Brody, D., Fisher, P.W., Bourdon, K., & Koretz, D.S. (2010). Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. *Pediatrics*, *125*, 75–81.

Mikolas, P., Tozzi, L., Doolin, K., Farrell, C., O'Keane, V., & Frodl, T. (2018). Hippocampal subfields volumes, endocrine stress axis and early life stress in major depressive disorder. *European Neuropsychopharmacology*, *28*, S71–S72.

Roberts, B.W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O.P. John, R.W. Robins, & L.A. Pervin (Eds.), *Handbook of personality. Theory and research* (3rd edn, pp. 375–398). New York: Guilford Publications.

Rohde, P., Lewinsohn, P.M., & Seeley, J.R. (1994). Are adolescents changed by an episode of major depression? *Journal of the American Academy of Child & Adolescent Psychiatry*, *33*, 1289–1298.

Thompson, R.A., Flood, M.F., & Goodvin, R. (2006). Social support and developmental psychopathology. *Developmental Psychopathology*, *3*, 1–37.

Weiss, B., & Garber, J. (2003). Developmental differences in the phenomenology of depression. *Development and Psychopathology*, *15*, 403–430.

Wichstrøm, L., Belsky, J., & Steinsbekk, S. (2017). Homotypic and heterotypic continuity of symptoms of psychiatric disorders from age 4 to 10 years: A dynamic panel model. *Journal of Child Psychology and Psychiatry*, *58*, 1239–1247.

Wichstrøm, L., Berg-Nielsen, T.S., Angold, A., Egger, H.L., Solheim, E., & Sveen, T.H. (2012). Prevalence of psychiatric disorders in preschoolers. *Journal of Child Psychology and Psychiatry*, *53*, 695–705.

Widaman, K.F., Ferrer, E., & Conger, R.D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives*, *4*, 10–18.

World Health Organization. (2019). Adolescent mental health. Available from: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health [last accessed 30 April 2020].

Wray, N.R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E.M., Abdellaoui, A., … & Sullivan, P.F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, *50*, 668.

Accepted for publication: 9 November 2020

Study/Paper II

Morken, I. S., Wichstrøm, L., Steinsbekk, S., & Viddal, K. R. Depression and Personality Across Adolescence—Within-person Analyses of a Birth Cohort. *Submitted.*

Study/Paper III

Morken, I. S., Viddal, K. R., Von Soest, T., & Wichstrøm, L. (2023). Explaining the Female Preponderance in Adolescent Depression—A Four-Wave Cohort Study. *Research on Child and Adolescent Psychopathology*, 1-11. https://doi.org/10.1007/s10802-023-01031-6

Research on Child and Adolescent Psychopathology
https://doi.org/10.1007/s10802-023-01031-6

Explaining the Female Preponderance in Adolescent Depression—A Four-Wave Cohort Study

Ida Sund Morken[1] · **Kristine Rensvik Viddal**[1] · **Tilmann von Soest**[1,2] · **Lars Wichstrøm**[1,3]

Accepted: 17 January 2023

Abstract
In the transition from childhood into adolescence, a female preponderance in depression emerges. Despite substantial empirical research to test theoretical propositions as to why this happens, our understanding is still limited. One explanation claims that girls become exposed to more stress (stress exposure model) whereas another proposes that girls become more vulnerable to the impact of stress (stress reactivity model) than boys when entering adolescence. Stressful life events (SLEs) and bullying victimization are established risk factors for adolescent depression. However, whether these factors contribute to the gender difference in depression is undetermined and thus investigated herein. Children (49.9% boys; $n = 748$) and parents from two birth cohorts in Trondheim, Norway, were followed biennially from ages 8 to 14 with clinical interviews about symptoms of depressive disorders and self-reports on SLEs. Teachers reported on bullying victimization. Prospective associations were investigated using an autoregressive latent trajectory model with structured residuals, examining within-person longitudinal associations while accounting for all time-invariant confounding effects. The number of depressive symptoms increased from ages 12 to 14 among girls. In the period before (ages 10 to 12), girls and boys were equally exposed to SLEs and bullying victimization. Increased stress (both SLEs and bullying victimization) at age 12 predicted increased depression at age 14 more strongly among girls than boys. Hence, increased impact—but not exposure—of SLEs and bullying victimization in girls may partly explain the emerging female preponderance in depression, in line with a stress reactivity model.

Keywords Major depressive disorder (MDD) · Dysthymia · Structural equation modelling · Negative life events · Sex difference

Depression is a common disorder worldwide, is often recurrent, and is among the leading causes of years lived with disabilities (GBD 2019 Mental Health Collaborators, 2022). Throughout most of the lifespan, depressive symptoms and disorders occur more frequently among women than men, and this female preponderance emerges in early adolescence, by at least age 12 (Salk et al., 2017). Two of the stressors that repeatedly have been identified as risk factors for child and adolescent depression are stressful life events (SLEs) (Ge et al., 1994, 2001), and bullying victimization (Christina et al., 2021). However, whether these stressors are involved in explaining the emerging female preponderance in depression needs further inquiry—a task we undertake herein.

With profoundness of the sudden gender difference in depression as a backdrop, several etiological models have been developed to account for this phenomenon (Cyranowski et al., 2000; Hankin & Abramson, 2001; Hankin et al., 2007; Hyde & Mezulis, 2020; Hyde et al., 2008; Nolen-Hoeksema & Girgus, 1994). A common element of these models was first proposed by Nolen-Hoeksema and Girgus (1994), namely that gender differences in stress exposure might lead to a female preponderance in depression. This potential mechanism has been termed the *stress exposure model* (e.g., Hammen, 2009b; Hankin et al., 2007), and posits that when girls approach adolescence they experience more stressors than boys. Examples of such stressors are sexual harassment (e.g., Skoog et al., 2016) and relational problems with peers and friends (for a review, see Rose & Rudolph, 2006). Moreover, according to a stress-generation hypothesis (Hammen, 2009a), depression may lead to characteristics and behaviors that increase interpersonal stress, and this process could be more

✉ Ida Sund Morken
ida.sund.morken@ntnu.no

[1] Department of Psychology, NTNU – Norwegian University of Science and Technology, Trondheim, Norway

[2] Department of Psychology, University of Oslo, Oslo, Norway

[3] Department of Child and Adolescent Psychiatry, St. Olavs University Hospital, Trondheim, Norway

Published online: 04 February 2023

pronounced in adolescent girls than boys (Hammen, 2009a). Regardless of how girls become exposed to more stress than boys, increased levels of stress may partly explain why they also become more depressed. The *stress reactivity model* (Hammen, 2009b; Hankin et al., 2007), in contrast, suggests that girls become more vulnerable to stress than boys when entering adolescence. According to this model, adolescent girls may experience heightened negative emotional reactivity and hence a stronger impact of stress on depression, which, in turn, may partly explain why more girls than boys become depressed. Naturally, these two explanations are not mutually exclusive (e.g., Hyde et al., 2008). To examine stress exposure and reactivity as explanations for the gender difference in adolescent depression, we propose five criteria—three for stress exposure and two for stress reactivity—that need to be fulfilled. Importantly, strong tests of the exposure and reactivity models involving SLEs and bullying victimization according to these criteria are lacking.

Stress Exposure Model

For the stress exposure model to be a valid explanation for the gender difference in depression, Criterion 1 states that girls should become exposed to more stress than boys just before the female preponderance in depression emerges (i.e., early adolescence), and not earlier (i.e., preadolescence). Previous research has found an increase in SLEs from childhood to adolescence (e.g., Larson & Ham, 1993). However, whether this increase is stronger for girls than boys is unclear. Even though a prior meta-analysis indicated that girls are exposed to more SLEs than boys, particularly during adolescence (Davis et al., 1999), later studies have not found gender differences in SLEs in adolescence (Jenness et al., 2019; Sund et al., 2003). Research on bullying victimization has also provided mixed results. Notably, most studies have investigated specific forms of bullying victimization, demonstrating, for example, that girls become more exposed to relational bullying and boys to physical bullying in mid-adolescence (Hager & Leadbeater, 2016), and that girls are more exposed to cyber victimization in early adolescence (Holfeld & Leadbeater, 2017). Other studies have not identified gender differences in the prevalence of relational bullying (Lepore & Kliewer, 2019) or cyber victimization (Díaz & Fite, 2019). However, given that all types of bullying victimization arguably thwart the fundamental need to belong, which in and by itself increases the risk for depression (Verhagen et al., 2018), we focus on bullying victimization in general. Studies on overall bullying victimization either portray an increase only among adolescent girls (Wendelborg, 2020), or find no such gender difference in prevalence (Sweeting et al., 2006). In sum, there is no consistent evidence indicating that girls become more exposed to SLEs or bullying victimization than boys just before the onset of the gender difference in depression.

Next, also pertaining to the stress exposure model, Criterion 2 states that increased stress in preadolescence should predict increased depression (at least among girls) in early adolescence, a prediction at the within-person level. Most prior research has utilized between-person information, asking whether those exposed to more SLEs and bullying victimization *than other adolescents* also become more depressed *than other adolescents*. However, other adolescents' stress exposure and level of depression cannot be involved in the development of depression. As advocated by several developmentalists (Berry & Willoughby, 2017; Hamaker et al., 2020), traditional analytical approaches, such as ordinary cross-lagged analyses of longitudinal data, do not disentangle within- from between-person information and therefore provide limited information from which to draw causal inferences (Berry & Willoughby, 2017). Importantly, the results from studies using between-person information can be influenced by time-invariant confounding effects, such as stable effects of genetics increasing the risk of both SLEs and depression (Clarke et al., 2018) or a persistent harsh parenting style increasing the risk for both bullying victimization and depression (Tang et al., 2018). Thus, to more closely approximate questions of causality (see e.g., Lervåg, 2020) while examining Criterion 2, we need to obtain information about the within-person association between preadolescent stress and prospective increased depression.

Finally, to support stress exposure as an explanation for the female preponderance in depression, Criterion 3 states that the sudden gender difference in depression should be accounted for (i.e., mediated) by an increasing gender disparity in levels of stress. However, at present, we do not know whether SLEs or bullying victimization mediate the gender difference in depression.

Stress Reactivity Model

To support a stress reactivity model, two criteria must be met. Criterion 4 states that increased stress should be more strongly associated with increased depression in girls than boys—and at the within-person level. Moreover, Criterion 5 states that this gender difference in stress reactivity should first appear in late childhood or early adolescence. Some regression-type research indicates that SLEs predict the level of depression to a stronger degree in girls than in boys during early adolescence (Ge et al., 1994), but whether this is the case for bullying victimization is unclear (Christina et al., 2021; Lepore & Kliewer, 2019). These findings notwithstanding, we lack research on whether associations are stronger among girls at the within-person level (Criterion 4) and whether they are specific to early adolescence as opposed to middle childhood (Criterion 5).

The Current Study

We investigate whether and how SLEs and bullying victimization contribute to explaining the emerging female preponderance in depression. We do this through biennial follow-ups of a community sample spanning from middle childhood to adolescence and by measuring depressive symptoms as defined by the Diagnostic and Statistical Manual of Mental Disorders (5th ed.; *DSM-5*; American Psychiatric Association, 2013). We examine whether the female preponderance in depression is partially explained by (i) increased stress exposure, where girls become more exposed to SLEs and bullying victimization than boys before this gender gap emerges, or (ii) increased stress reactivity, where SLEs and bullying victimization predict depressive symptoms to a stronger degree among girls than boys when entering adolescence. We hypothesize that both the stress exposure and reactivity explanations partly account for the emerging gender difference in depression. These explanations are examined adhering to the above-mentioned five criteria.

Methods

Participants and Procedure

The Trondheim Early Secure Study (TESS) (Steinsbekk & Wichstrøm, 2018) comprises children from the 2003 and 2004 birth cohorts in Trondheim, Norway ($N = 3{,}456$). A letter of invitation along with the Strengths and Difficulties Questionnaire (SDQ) version 4–16 (Goodman et al., 2000) was sent to the children's homes prior to the age 4 routine health check-up. Almost all children met with their parents at the check-up ($n = 3{,}358$). Parents received information about the study orally and in writing from the health nurse and written consent was obtained. Study procedures were approved by the Regional Committee for Medical and Health Research Ethics, Mid-Norway (approval number 2009/994).

To increase statistical power, children were divided into four strata based on their SDQ score (0–4, 5–8, 9–11, 12–40), and the probability of being selected increased with increasing scores (37%, 48%, 70%, and 89% from the respective strata). This oversampling of mental health problems was accounted for in the analyses. The drop-out rate after consent at the well-child clinic did not differ across the four SDQ strata; $\chi^2(3) = 5.70$, $p = 0.127$, or by gender; $\chi^2(3) = 0.23$, $p = 0.973$. Of the 1,250 children randomly selected for the study, 1,007 were successfully enrolled at Time 1 ($M_{age} = 4.59$, $SD = 0.25$; 49.1% boys) (for a flowchart, see Figure S1). Testing occurred biennially. Given that our research questions pertain to explaining depression in the transition from middle childhood to early adolescence, we included data from ages 8 (T3: $M_{age} = 8.79$, $SD = 0.23$), 10 (T4: $M_{age} = 10.51$, $SD = 0.17$), 12 (T5: $M_{age} = 12.50$, $SD = 0.14$) and 14 years (T6: $M_{age} = 14.35$, $SD = 0.14$). Attrition rates between waves of data collection were as follows: T3-T4: 0.14%, T4-T5: 5.26% and T5-T6: 4.51% (for more details, see Figure S1). Participants with information from at least one data wave composed the analytical sample ($n = 748$). Overall, attrition was unrelated to the study variables, including symptoms of Major Depressive Disorder (MDD) and dysthymia, and SLEs and bullying victimization measured at ages 4 and 6 years. However, more symptoms of MDD (OR = 1.39, 95% CI [1.15, 1.70]) and dysthymia (OR = 1.35, 95% CI [1.12, 1.64]) at age 12 predicted attrition at age 14, and more bullying victimization at age 6 predicted attrition at ages 10, 12 and 14 (all ORs = 1.02, 95% CI [1.01, 1.03]). The sample characteristics are presented in Table 1. Even though the above analyses suggested that attrition was selective according to study variables, they should be interpreted in the context of the number of attrition analyses conducted. An overall test, the Little Missing Completely at Random (MCAR) test (Little, 1988) was therefore conducted. The results showed that data were just barely shy of being MCAR, $\Delta\chi^2(220) = 256.01$, $p = 0.048$, whereas the normed Little's test was 1.16 (normed value < 2 suggesting missing at random–MAR) thus indicating that data were at least MAR.

Measures

Depressive symptoms were measured as symptoms of MDD and dysthymia according to DSM-5 criteria (American Psychiatric Association, 2013) using a semi-structured psychiatric interview, the Child and Adolescent Psychiatric Assessment (CAPA) (Angold & Costello, 2000). Children and parents were interviewed separately. A symptom was considered present if reported as occurring in the three months prior by either respondent. Inter-rater reliabilities among blinded coders of 15% of audiotapes of CAPA interviews were ICC = 0.87 for symptoms of MDD and ICC = 0.85 for symptoms of dysthymia. A symptom count score was created as the sum of MDD and dysthymia symptoms.

SLEs were measured by parent and child reports on 31 SLEs occurring since the last visit (two years), ranging from important life events (e.g., new sibling, parents separated or divorced) to very serious ones (e.g., sexual abuse) (see Appendix S1 for a complete list). A SLE was considered present if reported by either respondent, and a SLE total score was created as the sum of the number of SLEs. Given the wide range of seriousness in these events, we tested the possibility that any association between depression and life-events was driven by more serious or less serious events by comparing the correlations between depression and important life events to the correlations to events with

Table 1 Sample characteristics

Characteristics	%
Gender of child	
Male	48.9
Female	51.1
Gender of parent informant	
Male	16.7
Female	83.3
Parent informant	
Biological parent	98.3
Adoptive parent	1.3
Foster parent	0.4
Biological parents' marital status	
Married	59.3
Cohabitating > 6 months	21.9
Cohabitating < 6 months	0.4
Divorced/separated/no longer cohabitating	16.4
Widowed	0.1
Never lived together	1.9
Ethnic origin of biological mother	
Norwegian	93.0
Western Countries	2.7
Other Countries	4.3
Ethnic origin of biological father	
Norwegian	91.0
Western Countries	5.8
Other Countries	3.2
Informant parents' socioeconomic status	
Leader	17.5
Professional, higher level	30.1
Professional, lower level	30.1
Formally skilled worker	18.5
Farmer/fishermen	0.2
Unskilled worker	3.6

a substantial potential for grave physical and mental harm, using the Satorra-Bentler scaled chi-square difference test (Satorra & Bentler, 2001). Allowing these correlations to be different did not improve model fit as compared to the correlations being identical, $\Delta\chi^2(4) = 2.49$, $p = 0.952$, a fact suggesting that they did not differ and that the depression-SLE association was not different according to the seriousness of the SLEs.

Bullying victimization was measured by a teacher version of the Olweus Bully Victim Questionnaire (OBVQ) (Solberg & Olweus, 2003), completed by the participant's primary teacher. This teacher version of the OBVQ consists of five items pertaining to both physical bullying and social exclusion ($\alpha = 0.69$ to 0.79) tapping the frequency of physical harm, verbal abuse, social exclusion, been overlooked, and belongings hidden or destroyed, during the last 3 months. Response options ranges from *Never*, *Rarely*, *1–3 times per month*, *1–4 times per week* to *Everyday*.

Sociodemographic information on child and parent was reported by the parent during the diagnostic interview. Gender was coded (0 = boy, 1 = girl) based on the child's national identification number, in which the child's biological sex at birth is registered.

Statistical Analyses

As we did expect a change in the overall level of depression, and potentially also in SLEs and bullying victimization we employed autoregressive latent trajectory models with structured residuals (ALT-SR) (Berry & Willoughby, 2017) because they can accommodate linear and non-linear changes over time. Changes from one time point to the next was captured by latent change scores. In line with Orth et al.'s (2022) tentative suggestions, we considered cross-lagged associations with standardized regression coefficients of 0.03, 0.07, and 0.12 to indicate small, medium, and large effect sizes, respectively.

As our goal was to explain the female preponderance in depression, we focused on the age span when this gender difference is first expected to emerge—in early adolescence (i.e., ages 12 to 14) (Salk et al., 2017). To examine whether changes were specific to this age period we also analyzed the two age spans just prior to it (ages 8 to 10 and 10 to 12). Imbedded in the explanatory stress-exposure and stress-reactivity models is a causal relation between stress (e.g., bullying victimization) and depression. Hence, the increase in the exposure (stress) should occur before the increase in the outcome (depression). The possibility of a stress generating effect of depression (Hammen, 2009a) on SLEs and bullying victimization should be taken into account. Hence, a parallel increase in stress from ages 12 to 14 would not suffice as unequivocal predictor of change in depression from 12 to 14, because increased stress could be an effect of increased depression in the same period, not a predictor of it. Provided we find the expected increase in depression from ages 12 to 14, the increase in stress should therefore occur in period before, that is from ages 10 to 12.

The three criteria pertaining to the stress exposure model were tested in the following way: Criterion 1, whether girls become more exposed to stress than boys just prior to the emergence of a gender difference in depressive symptoms (i.e., ages 10 to 12) and not before (i.e., ages 8 to 10), was examined by inspecting whether the latent change in SLEs and bullying victimization increased among girls, specifically from ages 10 to 12 and not 8 to 10. Second, we examined whether any increases in SLEs or bullying victimization from ages 10 to 12 were predicted by female gender. Criterion 2, whether increased SLEs and bullying victimization predicted later depressive symptoms in girls at the within-person level, was tested by applying a modified version of the ALT-SR

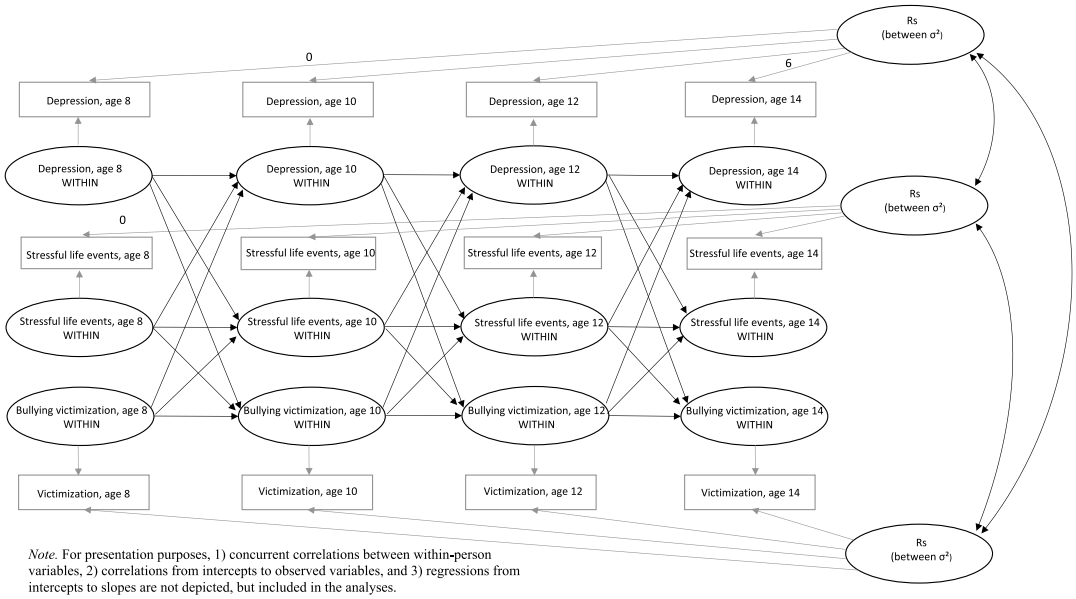
model (Berry & Willoughby, 2017) depicted in Fig. 1. In this model, between-person differences in depressive symptoms, SLEs and bullying victimization were captured by the intercept (representing the mean level) and slope (representing growth) in each of these three constructs, while within-person scores at each timepoint provide information about a person's deviation from his or her intercept and slope. In the traditional ALT-SR, the slope is set to be linear across all timepoints. Because the development of depressive symptoms, SLEs and bullying victimization are not necessarily expected to follow a linear pattern, we applied a latent basis model where the growth was freely estimated from the data, anchoring the slopes at ages 8 and 14. Criterion 3, whether the gender difference in depressive symptoms is explained by a potential increase in the study stressors, was examined by mediation analyses using Sobel's test (Mplus does not enable bootstrapping with population weights).

Regarding the stress reactivity explanation, Criterion 4, whether SLEs and bullying victimization predicted depressive symptoms more strongly for girls than boys at the within-person level, was tested by adding an interaction term between gender and SLEs and bullying victimization, respectively, at ages 10 and 12 in the ALT-SR models following procedures described by Mulder and Hamaker (2020). Finally, Criterion 5, whether a potentially stronger association for girls than boys was specific to early adolescence, was examined by inspecting whether the gender differences in the within-person associations between SLEs/bullying victimization and depressive symptoms were present only from ages 12 to 14 and not from ages 10 to 12.

All analyses were performed in Mplus 8.5 using a robust maximum likelihood estimator and probability weights to correct for the oversampling of children with mental health problems. Missing data were handled using a full information maximum likelihood (FIML) procedure under the assumption that data was MAR.

Results

The results showed rather low counts of depressive symptoms at ages 8 to 12, with scores between 1.0 and 1.5 for both genders, but with a sudden increase for 14-year-old girls to 2.1, while boys' depressive symptoms count remained at a stable low level (see Table 2 and Fig. 2). This increase was mirrored by female gender being associated with depressive symptoms at age 14 ($r = 0.16$, 95% CI [0.05, 0.24]) but not at ages 12 ($r = 0.08$, 95% CI [-0.03, 0.17]) or 10 ($r = 0.02$, 95% CI [-0.07, 0.10]). At age 8, female gender was associated with fewer symptoms ($r = -0.13$, 95% CI



Note. For presentation purposes, 1) concurrent correlations between within-person variables, 2) correlations from intercepts to observed variables, and 3) regressions from intercepts to slopes are not depicted, but included in the analyses.

Fig. 1 Theoretical autoregressive latent trajectory model with structural residuals model of depression, stressful life events and bullying victimization

Table 2 Gender differences in stressful life events (SLEs) and bullying victimization, and mean level change, ages 8–14

	Boys		Girls		Gender differences in latent change scores
	Mean level (SD)	*p*-value of 2-year change	Mean level (SD)	*p*-value of 2-year change	B [95% CI]
Depression					
Age 8	1.26 (1.77)		1.02 (1.66)	-	-
Age 10	1.13 (1.65)	0.176	1.31 (1.79)	0.053	0.14 [-0.02, 0.29]
Age 12	1.11 (1.96)	0.874	1.54 (2.04)	0.214	0.17 [-0.02, 0.35]
Age 14	1.16 (2.09)	0.061	2.05 (2.89)	0.003	0.23 [0.01, 0.44]
Stressful life events					
Age 8	1.10 (1.35)		0.96 (1.33)	-	-
Age 10	1.14 (1.35)	0.724	1.17 (1.51)	0.038	0.03 [-0.07, 0.13]
Age 12	1.63 (1.81)	< 0.001	1.63 (1.67)	< 0.001	-0.00 [-0.12, 0.11]
Age 14	1.94 (1.66)	0.047	2.20 (1.79)	< 0.001	0.11 [-0.01, 0.23]
Bullying victimization					
Age 8	1.08 (1.04)		1.03 (1.10)		
Age 10	1.02 (1.08)	0.349	0.89 (0.96)	0.175	-0.04 [-0.11, 0.03]
Age 12	0.99 (1.41)	0.854	0.85 (0.98)	0.954	-0.04 [-0.12, 0.04]
Age 14	0.60 (0.90)	0.001	0.66 (0.90)	0.002	0.03 [-0.04, 0.09]

[-0.23, -0.05]). Latent change scores analyses confirmed the emerging female preponderance in depression in early adolescence, as female gender predicted an increased number of depressive symptoms from ages 12 to 14 but not from ages 10 to 12 or 8 to 10 (see Table 2). Correlations between study variables are provided in Table S1.

Stress Exposure Model

To examine the stress exposure explanation, we first tested Criterion 1—girls becoming more exposed to stress than boys in the period prior to the female preponderance in depressive symptoms. In girls, SLEs significantly increased from ages

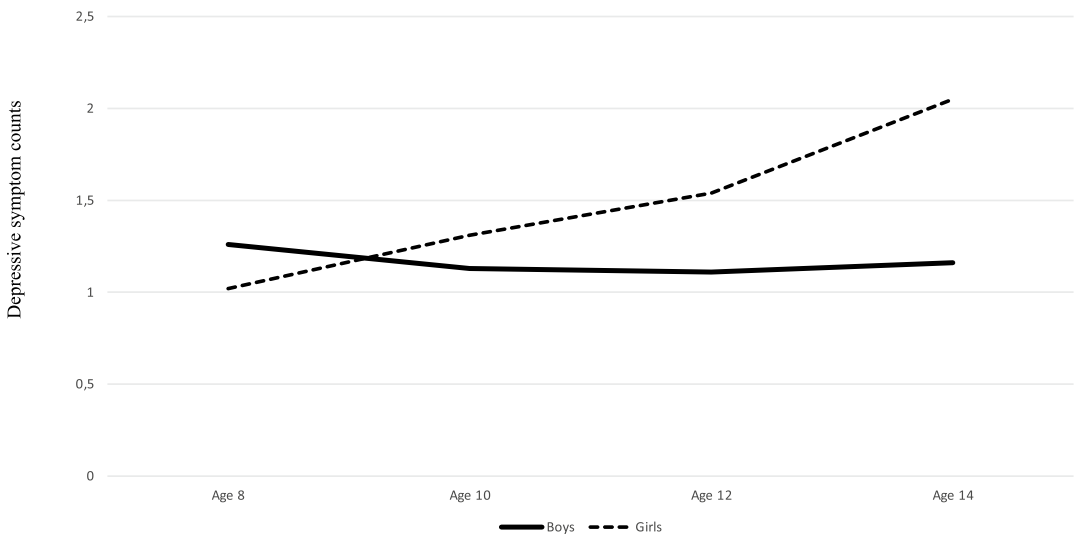


Fig. 2 Development of depressive symptom counts from age 8 to age 14 for boys and girls

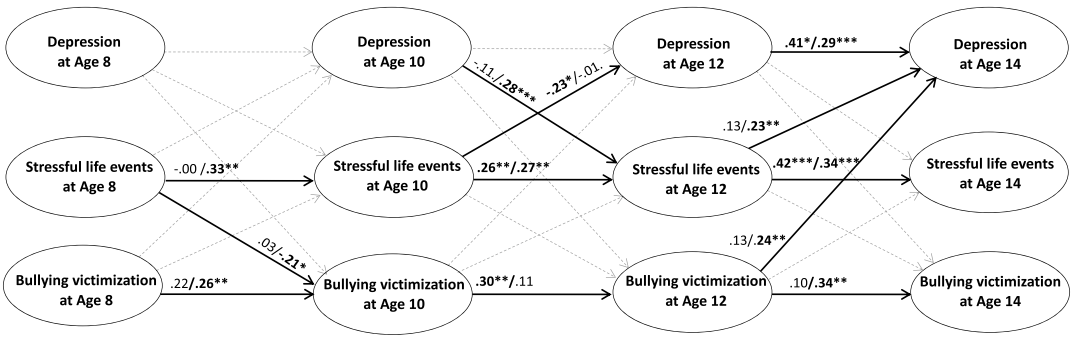
8 to 10 and 10 to 12, while bullying victimization remained stable (see Table 2). However, girls did not become more exposed to either SLEs or bullying victimization than boys in any of these age spans (see Table 2). Thus, Criterion 1 was not fulfilled. Regarding Criterion 2—stress predicting an increased number of depressive symptoms in girls at the within-person level, girls' depressive symptoms at age 14 were predicted by SLEs (B = 0.39, *SE* = 0.12, 95% CI [0.15, 0.62]) and bullying victimization (B = 0.67, *SE* = 0.25, 95% CI [0.18, 1.15]) at age 12 in ALT-SR analyses. Standardized regression coefficients of the associations are presented in Fig. 3 and indicate large effect sizes between the associations of both SLEs and bullying with girls' depressive symptoms at age 14 (i.e., $\beta \geq 0.12$). However, when examining Criterion 3—the gender effect on depressive symptoms at age 14 being mediated by increased SLEs or bullying victimization–at age 12, the results revealed no such effects (B = -0.01, *SE* = 0.03, 95% CI [-0.08, 0.05] and B = -0.04, *SE* = 0.04, 95% CI [-0.12, 0.03], respectively). In sum, Criteria 1 and 3 for increased stress exposure were not met.

Stress Reactivity Model

When examining the stress reactivity model, Criterion 4—stress predicting depressive symptoms more strongly in girls than boys at the within-person level—SLEs at age 12 predicted depressive symptoms at age 14 among girls (B = 0.39, *SE* = 0.12, 95% CI [0.15, 0.62]) but not among boys (B = 0.16, *SE* = 0.12, 95% CI [-0.08, 0.40]) (see Fig. 3 for standardized values), and a significant gender*SLEs interaction (B = 0.21, *SE* = 0.10, 95% CI [0.02, 0.39]) was detected. A similar pattern was observed for bullying victimization, which predicted depressive symptoms in girls (B = 0.67, *SE* = 0.25, 95% CI [0.18, 1.15]) but not boys (B = 0.25, *SE* = 0.21, 95% CI [-0.16, 0.66]) (see Fig. 3 for standardized values), with a significant gender*bullying victimization interaction (B = 0.53, *SE* = 0.20, 95% CI [0.15, 0.92]). Finally, Criterion 5 states that this gender difference should not be present before the emerging female preponderance in depression. As seen in Fig. 3, no such predictive effects were detected at younger ages, a finding supported by a nonsignificant gender*SLEs interaction (B = -0.06, *SE* = 0.06, 95% CI [-0.18, 0.06]) and gender*bullying victimization interaction (B = 0.01, *SE* = 0.14, 95% CI [-0.27, 0.29]) from ages 10 to 12. Thus, both criteria pertaining to the stress reactivity explanation were met.

Discussion

The female preponderance in depression first emerges in early adolescence, and adolescent SLEs and bullying victimization are consistent predictors of depression. However, it is undetermined whether SLEs and bullying victimization can explain why the gender difference in depression emerges at this point in development. We examined two psychosocial explanations—a stress exposure model (girls becoming more exposed to stress than boys) and a stress reactivity model (early adolescent girls reacting to stressors with more depression than boys) by systematically testing whether specific criteria were fulfilled. We examined these criteria by drawing on a representative birth cohort sample followed biennially from age 8 to 14 years and applying ALT-SR methodology to illuminate within-person development. Our results showed the expected female preponderance in depression: the number of DSM-5 defined symptoms of major depression and dysthymia increased sharply from ages 12 to 14 among girls but not among boys. Furthermore, a stress exposure explanation



Note. Bold arrows represent statistically significant predictions for boys/girls, or both.

Fig. 3 Within-person standardized regression coefficients from autoregressive latent trajectory model with structured residuals for boys/girls

was not supported, whereas a stress reactivity explanation was supported; girls who were exposed to more SLEs and bullying victimization at age 12 developed an increased number of depressive symptoms at age 14, with standardized regression coefficients indicating large effect sizes. No such associations were seen among boys or at earlier timepoints. Notably, the five criteria proposed herein may be applied in future studies on other stressors' role in explaining the female preponderance in depression.

Although the present study was not positioned to unravel the underlying mechanisms for the assumed stress reactivity, we draw attention to some possibilities. First, puberty likely plays a role in stress regulation. Gonadal hormone secretion increases in puberty, which, in turn, is associated with gender differences in the hypothalamic–pituitary–adrenal axis response to stress, including cortisol production (Heck & Handa, 2019). Of current interest, there is evidence that adolescent girls evince lower cortisol levels when exposed to social stress than boys (Bouma et al., 2009), which may increase their vulnerability to developing depression (Colich et al., 2015). Arguably, therefore, the altered stress regulation in puberty may reinforce the negative effects of SLEs and bullying victimization in girls in particular. Another potential mechanism involves adolescent girls' use of maladaptive cognitive coping strategies when faced with stress, such as rumination, which increases the risk for depression (Aldao et al., 2010). Indeed, girls tend to ruminate more than boys and more so in early adolescence than in late childhood (Hampel & Petermann, 2005), making them more vulnerable to depressive reactions to stress in this particular developmental period. To clarify the practical implications of the present findings, future studies should delineate the mechanisms involved in girls' increased reactivity to SLEs and bullying victimization. In turn, preventive and treatment efforts may target the most potent mechanisms.

Our study found that girls did not become more exposed to either SLEs or bullying victimization than boys from age 10 to age 12, and these stressors did not mediate the gender effect on depressive symptoms. In effect, the increased stress exposure model was not supported. Previous studies have reported mixed evidence for gender differences in the prevalence of or increase in SLEs and bullying victimization in early adolescence, and the current findings coincide with those reporting no gender difference in SLEs (Jenness et al., 2019; Sund et al., 2003) or overall bullying victimization (Sweeting et al., 2006) in early adolescence. The discrepancies between findings may be attributed to a range of methodological and sample differences, including differences in the specific SLEs studied, age of participants, secular period, populations, and nationalities. For example, Hankin et al. (2007) found in a sample of US adolescents that girls were exposed to more interpersonal stress than the boys were, whereas Sund et al. (2003) found that Norwegian adolescent girls and boys were exposed to a similar amount of interpersonal stress.

We focused on two stressors that are established risk factors for depression: SLEs and bullying victimization. Whether other relevant stressors, such as daily hassles (Hankin et al., 2007) and peer sexual harassment victimization (Dahlqvist et al., 2016), follow exposure or reactivity patterns awaits future research. Importantly, to provide strong tests of exposure and reactivity explanations, such inquiries should cover the whole age-span from late childhood (i.e., even *before* the gender difference in depression appears) until adolescence. In a related vein, SLEs and victimization may have different effects on maintaining or widening the gender difference in depression in later adolescence (Salk et al., 2017), and other contributing factors may differ between these developmental periods. Our results are therefore specific to the development of depressive symptoms in the early adolescent period. Finally, the present results are specific to symptoms of depression and do not preclude the possibility of boys reacting more strongly than girls with symptoms of other disorders.

At age 8, boys evinced slightly more depressive symptoms than girls. As pointed out in previous research (reviewed by e.g., Salk et al., 2017), there are some reports of a male preponderance in depression in early childhood, whereas others do not find this difference. Hence, whether there is a gender difference in early childhood awaits further inquiry.

Limitations

While this study has a range of strengths, including a representative community sample, clinical interviews to assess depressive symptoms, multiple informants, repeated assessments before and through the crucial years when the gender differences emerge, and a solid statistical approach to assess predictions at the within-person level, we acknowledge several limitations. First, children with more depressive symptoms at age 12 more often dropped out of the study by age 14, potentially resulting in underestimating the increase in depressive symptoms during this period. However, considering that our prime interest was gender differences in prevalence and associations and that we applied an FIML approach to missingness, selective attrition likely did not have a major impact on the results. Second, although we adjusted for all time-invariant confounders, time-varying factors, such as bodily changes associated with puberty or increased risk behavior in adolescence, may still have influenced both stress and depressive symptoms. Third, we studied symptom counts. Although there is no compelling evidence pointing to depression being categorical in nature (Haslam et al., 2012), we cannot be sure that our findings apply to depressive disorders. Fourth, we captured depressive symptoms occurring in the prior 3-month period, and symptoms occurring

between our 2-year intervals of observation might have been missed. Fifth, we summed the number of SLEs, which differed considerably in frequency and seriousness; thus, we were not able to discern the effect of specific SLEs. Sixth, we assessed bullying victimization based on teacher reports. As such, bullying victimization that occurs outside of the school context, perhaps most notably cyber victimization (Díaz & Fite, 2019), might not have been captured. Even though those who are victimized online are often victimized at school as well (Wendelborg, 2020), the rate of overall bullying victimization might have been deflated. Seventh, gender was measured as biological sex assigned by birth (either girl or boy), thus not taking gender identity into account. Mounting evidence suggests that non-binary youth are at increased risk for psychiatric symptoms (e.g., Johansson et al., 2022; Price-Feeney et al., 2020) and theories on the gender difference in depression are limited by the gender binary. Notably, current surveys in Norway indicate that 0.4% of adults do not consider themselves as males or females and that 0.005% do not know (Statistics Norway, 2021). Although these numbers likely are higher among youths, the models that were tested herein would demand larger sample size than ours. Finally, future studies should include direct assessments of stress reactivity (e.g., behavioral observation or electrophysiological or hormonal measures).

Conclusions

The current study is the first to examine and present support for the notion that increased reactivity to both SLEs and bullying victimization in early adolescent girls may contribute to explaining the emerging female preponderance in depression. These findings highlight the transition to early adolescence as critical for preventive interventions. Professionals implementing such efforts should take into account that exposure to SLEs and bulling victimization, occurring already in preadolescence, might confer a heightened risk for depressive symptoms for early adolescent girls in particular.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10802-023-01031-6.

Acknowledgements We are grateful to the adolescents and families who diligently biennially participated in the Trondheim Early Secure Study (TESS).

Authors' Contribution Statements All authors conceived the research questions and design. ISM performed the literature searches and drafted the first manuscript. LW, KRV and TVS contributed to the development and writing of the manuscript. Statistical analyses were performed by LW, and statistical interpretation was performed by all authors. All authors approved the final draft. All authors had full access to all data in the study and final responsibility for the decision to submit for publication.

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl St. Olavs Hospital - Trondheim University Hospital). The paper was funded by grant ES611813 from the Research Council of Norway, and a grant from the Liaison Committee for Education, Research and Innovation in Central Norway. TVS's work with this article was supported grant no. 288083 and no. 301010 from the Research Council of Norway.

Materials and/or Code Availability Due to conditions for consent from participants, data cannot be shared.

Compliance With Ethical Standards

Conflict of Interest We declare no competing interests. The funders of the study had no role in the study design, data collection, data analyses, data interpretation, or writing of the report.

Ethics Approval Study procedures were approved by the Regional Committee for Medical and Health Research Ethics, Mid-Norway (approval number 2009/994).

Inform Consent We have obtained written consent from the participants.

References

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review, 30*(2), 217–237. https://doi.org/10.1016/j.cpr.2009.11.004

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Angold, A., & Costello, E. J. (2000). The child and adolescent psychiatric assessment (CAPA). *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(1), 39–48. https://doi.org/10.1017/S003329170003498X

Berry, D., & Willoughby, M. T. (2017). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child Development, 88*(4), 1186–1206. https://doi.org/10.1111/cdev.12660

Bouma, E. M., Riese, H., Ormel, J., Verhulst, F. C., & Oldehinkel, A. J. (2009). Adolescents' cortisol responses to awakening and social stress; effects of gender, menstrual phase and oral contraceptives. *The TRAILS Study. Psychoneuroendocrinology, 34*(6), 884–893. https://doi.org/10.1016/j.psyneuen.2009.01.003

Christina, S., Magson, N. R., Kakar, V., & Rapee, R. M. (2021). The bidirectional relationships between peer victimization and internalizing problems in school-aged children: An updated systematic review and meta-analysis. *Clinical Psychology Review, 85*, 1–19. https://doi.org/10.1016/j.cpr.2021.101979

Clarke, T. -K., Zeng, Y., Navrady, L., Xia, C., Haley, C., Campbell, A., Navarro, P., Amador, C., Adams, M. J., Howard, D. M., Soler, A.,

Hayward, C., Thomson, P. A., Smith, B. H., Padmanabhan, S., Hocking, L. J., Hall, L. S., Porteus, D. J., Deary, I. J., & McIntosh, A. M. (2018). Genetic and environmental determinants of stressful life events and their overlap with depression and neuroticism. *Wellcome Open Research, 3*(11), 1–17. https://doi.org/10.12688/wellcomeopenres.13893.2

Colich, N. L., Kircanski, K., Foland-Ross, L. C., & Gotlib, I. H. (2015). HPA-axis reactivity interacts with stage of pubertal development to predict the onset of depression. *Psychoneuroendocrinology, 55*, 94–101. https://doi.org/10.1016/j.psyneuen.2015.02.004

Cyranowski, J. M., Frank, E., Young, E., & Shear, M. K. (2000). Adolescent onset of the gender difference in lifetime rates of major depression: A theoretical model. *Archives of General Psychiatry, 57*(1), 21–27. https://doi.org/10.1001/archpsyc.57.1.21

Dahlqvist, H. Z., Landstedt, E., Young, R., & Gådin, K. G. (2016). Dimensions of peer sexual harassment victimization and depressive symptoms in adolescence: A longitudinal cross-lagged study in a Swedish sample. *Journal of Youth and Adolescence, 45*(5), 858–873. https://doi.org/10.1007/s10964-016-0446-x

Davis, M. C., Matthews, K. A., & Twamley, E. W. (1999). Is life more difficult on Mars or Venus? A meta-analytic review of sex differences in major and minor life events. *Annals of Behavioral Medicine, 21*(1), 83–97. https://doi.org/10.1007/BF02895038

Díaz, K. I., & Fite, P. J. (2019). Cyber victimization and its association with substance use, anxiety, and depression symptoms among middle school youth. *Child & Youth Care Forum, 48*(4), 529–544. https://doi.org/10.1007/s10566-019-09493-w

Ge, X., Conger, R. D., & Elder, G. H., Jr. (2001). Pubertal transition, stressful life events, and the emergence of gender differences in adolescent depressive symptoms. *Developmental Psychology, 37*(3), 404–417. https://doi.org/10.1037/0012-1649.37.3.404

Ge, X., Lorenz, F. O., Conger, R. D., Elder, G. H., & Simons, R. L. (1994). Trajectories of stressful life events and depressive symptoms during adolescence. *Developmental Psychology, 30*(4), 467–483. https://doi.org/10.1037/0012-1649.30.4.467

GBD 2019 Mental Disorders Collaborators. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Psychiatry, 9*(2), 137–150. https://doi.org/10.1016/S2215-0366(21)00395-3

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry, 177*(6), 534–539. https://doi.org/10.1192/bjp.177.6.534

Hager, A. D., & Leadbeater, B. J. (2016). The longitudinal effects of peer victimization on physical health from adolescence to young adulthood. *Journal of Adolescent Health, 58*(3), 330–336. https://doi.org/10.1016/j.jadohealth.2015.10.014

Hamaker, E. L., Mulder, J. D., & van IJzendoorn, M. H. (2020). Description, prediction and causation: Methodological challenges of studying child and adolescent development. *Developmental Cognitive Neuroscience, 46*, 1–14. https://doi.org/10.1016/j.dcn.2020.100867

Hammen, C. (2009a). Adolescent depression: Stressful interpersonal contexts and risk for recurrence. *Current Directions in Psychological Science, 18*(4), 200–204. https://doi.org/10.1111/j.1467-8721.2009.0163

Hammen, C. (2009b). Stress exposure and stress generation in adolescent depression. In S. Nolen-Hoeksema & L. M. Hilt (Eds.), *Handbook of depression in adolescents* (pp. 305–333). Routledge/Taylor & Francis Group.

Hampel, P., & Petermann, F. (2005). Age and gender effects on coping in children and adolescents. *Journal of Youth and Adolescence, 34*(2), 73–83. https://doi.org/10.1007/s10964-005-3207-9

Hankin, B. L., & Abramson, L. Y. (2001). Development of gender differences in depression: An elaborated cognitive vulnerability–transactional stress theory. *Psychological Bulletin, 127*(6), 773–796. https://doi.org/10.1037/0033-2909.127.6.773

Hankin, B. L., Mermelstein, R., & Roesch, L. (2007). Sex differences in adolescent depression: Stress exposure and reactivity models. *Child Development, 78*(1), 279–295. https://doi.org/10.1111/j.1467-8624.2007.00997.x

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: A quantitative review of taxometric research. *Psychological Medicine, 42*(5), 903–920. https://doi.org/10.1017/S0033291711001966

Heck, A. L., & Handa, R. J. (2019). Sex differences in the hypothalamic–pituitary–adrenal axis' response to stress: An important role for gonadal hormones. *Neuropsychopharmacology, 44*(1), 45–58. https://doi.org/10.1038/s41386-018-0167-9

Holfeld, B., & Leadbeater, B. J. (2017). Concurrent and longitudinal associations between early adolescents' experiences of school climate and cyber victimization. *Computers in Human Behavior, 76*, 321–328. https://doi.org/10.1016/j.chb.2017.07.037

Hyde, J. S., & Mezulis, A. H. (2020). Gender differences in depression: Biological, affective, cognitive, and sociocultural factors. *Harvard Review of Psychiatry, 28*(1), 4–13. https://doi.org/10.1097/HRP.0000000000000230

Hyde, J. S., Mezulis, A. H., & Abramson, L. Y. (2008). The ABCs of depression: Integrating affective, biological, and cognitive models to explain the emergence of the gender difference in depression. *Psychological Review, 115*(2), 291–331. https://doi.org/10.1037/0033-295X.115.2.291

Jenness, J. L., Peverill, M., King, K. M., Hankin, B. L., & McLaughlin, K. A. (2019). Dynamic associations between stressful life events and adolescent internalizing psychopathology in a multiwave longitudinal study. *Journal of Abnormal Psychology, 128*(6), 596–609. https://doi.org/10.1037/abn0000450

Johansson, C., Kullgren, C., Bador, K., & Kerekes, N. (2022). Gender non-binary adolescents' somatic and mental health throughout 2020. *Frontiers in Psychology, 13*, 1–14. https://doi.org/10.3389/fpsyg.2022.993568

Larson, R., & Ham, M. (1993). Stress and" storm and stress" in early adolescence: The relationship of negative events with dysphoric affect. *Developmental Psychology, 29*(1), 130–140. https://doi.org/10.1037/0012-1649.29.1.130

Lepore, S. J., & Kliewer, W. (2019). Social intelligence attenuates association between peer victimization and depressive symptoms among adolescents. *Psychology of Violence, 9*(6), 644–652. https://doi.org/10.1037/vio0000234

Lervåg, A. (2020). Some roads less travelled—different routes to understanding the causes of child psychopathology. *Journal of Child Psychology and Psychiatry, 61*(6), 625–627. https://doi.org/10.1111/jcpp.13274

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198–1202. https://doi.org/10.2307/2290157

Mulder, J. D., & Hamaker, E. L. (2020). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(4), 268–468. https://doi.org/10.1080/10705511.2020.1784738

Nolen-Hoeksema, S., & Girgus, J. S. (1994). The emergence of gender differences in depression during adolescence. *Psychological Bulletin, 115*(3), 424–443. https://doi.org/10.1037/0033-2909.115.3.424

Orth, U., Meier, L. L., Bühler, J. L., Dapp, L. C., Krauss, S., Messerli, D., & Robins, R. W. (2022). Effect Size Guidelines for Cross-Lagged Effects. *Psychological Methods*. https://doi.org/10.1037/met0000499

Price-Feeney, M., Green, A. E., & Dorison, S. (2020). Understanding the mental health of transgender and nonbinary youth. *Journal*

of Adolescent Health, 66(6), 684–690. https://doi.org/10.1016/j.jadohealth.2019.11.314

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin, 132*(1), 98–131. https://doi.org/10.1037/0033-2909.132.1.98

Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin, 143*(8), 783–822. https://doi.org/10.1037/bul0000102

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507–514. https://doi.org/10.1007/BF02296192

Skoog, T., Özdemir, S. B., & Stattin, H. (2016). Understanding the link between pubertal timing in girls and the development of depressive symptoms: The role of sexual harassment. *Journal of Youth and Adolescence, 45*(2), 316–327. https://doi.org/10.1007/s10964-015-0292-2

Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression, 29*(3), 239–268. https://doi.org/10.1002/ab.10047

Statistics Norway (SSB). (2021). *Kjønn, identitet og seksualitet i kvantitative spørreundersøkelser. Veileder for spørsmål om seksuell orientering, kjønnsidentitet, kjønnsuttrykk og kjønnskarakteristika, 2021/28.* Statistics Norway. Retrieved November 15, 2022, from https://www.ssb.no/sosiale-forhold-og-kriminalitet/levekar/artikler/kjonn-identitet-og-seksualitet-i-kvantitative-sporreundersokelser/_/attachment/inline/86aeca7c-c6f5-4eee-a00a-8274c5261bb5:5c0f2c2442265393def722af81eca0530dba4539/NOT2021-28_web.pdf

Steinsbekk, S., & Wichstrøm, L. (2018). Cohort Profile: The Trondheim Early Secure Study (TESS)—a study of mental health, psychosocial development and health behaviour from preschool to adolescence. *International Journal of Epidemiology, 47*(5), 1401–1401. https://doi.org/10.1093/ije/dyy190

Sund, A. M., Larsson, B., & Wichstrøm, L. (2003). Psychosocial correlates of depressive symptoms among 12–14-year-old Norwegian adolescents. *Journal of Child Psychology and Psychiatry, 44*(4), 588–597. https://doi.org/10.1111/1469-7610.00147

Sweeting, H., Young, R., West, P., & Der, G. (2006). Peer victimization and depression in early–mid adolescence: A longitudinal study. *British Journal of Educational Psychology, 76*(3), 577–594. https://doi.org/10.1348/000709905X49890

Tang, A. -M., Deng, X. -L., Du, X. -X., & Wang, M. -Z. (2018). Harsh parenting and adolescent depression: Mediation by negative self-cognition and moderation by peer acceptance. *School Psychology International, 39*(1), 22–37. https://doi.org/10.1177/0143034317709066

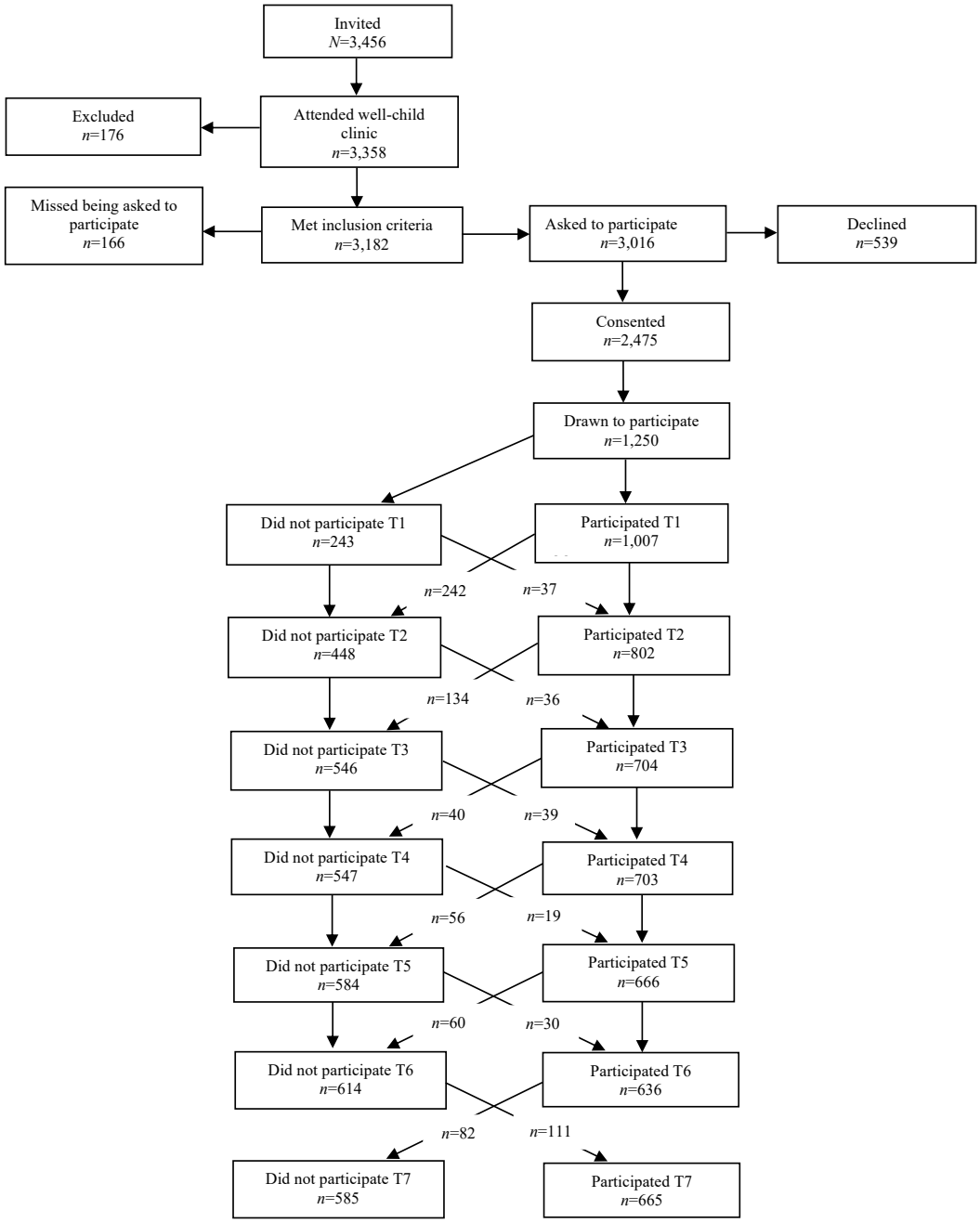
Verhagen, M., Lodder, G. M., & Baumeister, R. F. (2018). Unmet belongingness needs but not high belongingness needs alone predict adverse well-being: A response surface modeling approach. *Journal of Personality, 86*(3), 498–507. https://doi.org/10.1111/jopy.12331

Wendelborg, C. (2020). *Mobbing og arbeidsro i skolen: Analyse av Elevundersøkelsen skoleåret 2019/20.* NTNU Samfunnsforskning AS. http://hdl.handle.net/11250/2639017

Appendix

Figure I

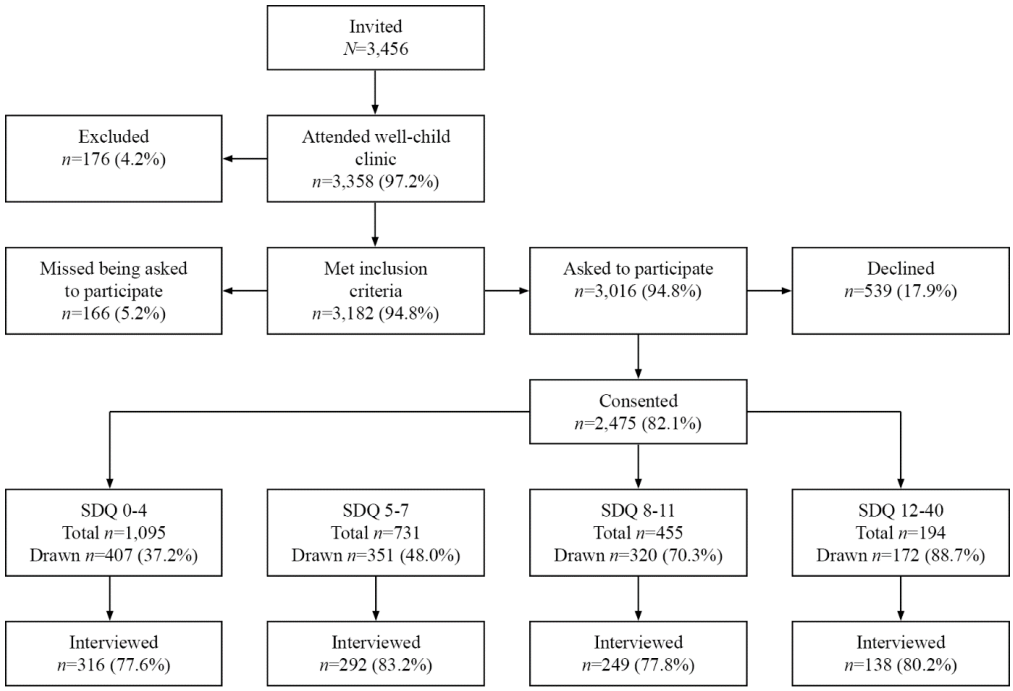
Flow chart of recruitment and follow-up



Note. Number of participants at the various assessment points is based on the number of participants invited to participate (n = 1,250) minus those who did not participate at the respective measurement point (i.e., T1, T2).

Figure II

TESS Stratification



Note. SDQ=Strengths and Difficulties Questionnaire (Goodman et al., 2000).

Supplementary Tables Study I (S1-S6)

Table S1

Stability relative to others (Pearson's correlation coefficient) stability relative to others (Intra-class correlation coefficient) between number of symptoms of MDD and Dysthymia, age 4-14

		1	2	3	4	5	6	7	8	9	10	11	12
1.	MDD – Age 4		.88***	.17***	.16**	.29***	.27***	.08	.10*	.09	.06	.07	.07
2.	Dysthymia – Age 4			.13**	.15**	.24***	.23***	.04	.08	.08	.04	.08	.08
3.	MDD – Age 6	*ICC* = .22 [.16, .27]			.87***	.35***	.32***	.30***	.29***	.31***	.31***	.27***	.29***
4.	Dysthymia – Age 6		*ICC* = .22 [.16, .26]			.36***	.33***	.24***	.25***	.26***	.26***	.20***	.22***
5.	MDD – Age 8	*ICC* = .28 [.22, .32]		*ICC* = .29 [.23, 34]			.89***	.31***	.30***	.31***	.37***	.24***	.24***
6.	Dysthymia – Age 8		*ICC* = .21 [.16, .22]		*ICC* = .31 [.24, .35]			.32***	.33***	.33***	.37***	.25***	.25***

Table S1

Stability relative to others (Pearson's correlation coefficient) stability relative to others (Intra-class correlation coefficient) between number of symptoms of MDD and Dysthymia, age 4-14

7.	MDD – Age 10	*ICC* = .12 [-.19 .26]		*ICC* = .22 [.17 .25]		*ICC* = .26 [.22, .29]		.90***	.35***	.37***	.35***	.35***
8.	Dysthymia – Age 10		*ICC* = .08 [-.11, .18]		*ICC* = .19 [.12, .23]		*ICC* = .32 [.26, .35]		.37***	.39***	.35***	.36***
9.	MDD – Age 12	*ICC* = .09 [.01 .13]		*ICC* = .19 [.14 .22]		*ICC* = .26 [.21 .28]	*ICC* = .32 [.32, .33]			.91***	.55***	.54***
10.	Dysthymia – Age 12		*ICC* = .03 [.05, .08]		*ICC* = .17 [.19, .22]		*ICC* = .28 [.23, .31]	*ICC* = .35 [.32, .36]			.52***	.53***
11.	MDD – Age 14	*ICC* = .05 [-.03 .10]		*ICC* = .12 [.07 .16]		*ICC* = .18 [.11 .22]	*ICC* = .28 [.22 .32]		*ICC* = .47 [.43, .48]			.93***
12.	Dysthymia – Age 14		*ICC* = .02 [-.05, .07]		*ICC* = .07 [.01, .11]		*ICC* = .18 [.10, .23]	*ICC* = .28 [.22, .32]		*ICC* = .46 [.43, .48]		

Note: *r* = Pearson Product-Number Coefficient, *ICC*=Intraclass Correlation Coefficient, * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [95% CI for ICC]

Table S2

Percentages with symptoms at age 6 sorted by symptoms at age 4

	Age 6: number of symptoms					
Age 4: number of symptoms	0	1	2	3	4	5
0	70.9%	21.2%	4.9%	2.2%	0.6%	0.2%
1	56.7%	25.8%	11.9%	3.9%	1.7%	0%
2	47.7%	31.1%	17.3%	1.6%	1.8%	0.5%
3	32.1%	37.0%	12.2%	0%	18.7%	0%
4	41.4%	49.8%	8.8%	0%	0.0%	0%
5	0%	0%	0%	0%	100%	0%

Table S3

Percentages with symptoms at age 8 sorted by symptoms at age 6

	Age 8: number of symptoms					
Age 6: number of symptoms	0	1	2	3	4	5
0	72.4%	22.2%	3.9%	1.1%	0.3%	0%
1	65.0%	25.1%	7.2%	2.0%	0.6%	0%
2	54.9%	20.7%	11.3%	8.2%	3.0%	2.0%
3	34.1%	34.0%	9.8%	20.5%	1.6%	0%
4	31.9%	26.5%	37.6%	0%	4.0%	0%
5	67.9%	23.1%	5.9%	2.4%	0.7%	0.2%

Table S4

Percentages with symptoms at age 10 sorted by symptoms at age 8

	Age 10: number of symptoms						
Age 8: number of symptoms	0	1	2	3	4	5	6
0	70.8%	22.0%	5.2%	1.3%	0.4%	0.3%	0.0%
1	51.6%	35.3%	6.1%	3.6%	1,4%	1.0%	1.0%
2	50.9%	22.9%	18.9%	6.6%	0.7%	0%	0%
3	26.1%	15.0%	41.7%	13.5%	3.7%	0%	0%
4	33.1%	28.8%	33.1%	5.9%	0%	0%	0%
5	0%	100%	0%	0%	0%	0%	0%
6	63.7%	25.1%	7.4%	2.5%	0.7%	0.5%	0.2%

Table S5

Percentages with symptoms at age 12 sorted by symptoms at age 10

	Age 12: number of symptoms							
Age 10: number of symptoms	0	1	2	3	4	5	6	7
0	74.1%	20.9%	3.7%	0.5%	0.6%	0.2%	0%	0%
1	50.2%	29.0%	10.6%	4.3%	2.6%	1.4%	0.9%	0.9%
2	34.6%	22.9%	32.7%	5.7%	4.1%	0%	0%	0%
3	25.0%	18.4%	27.5%	12.5%	16.6%	0%	0%	0%
4	13.4%	27.3%	13.4%	45.8%	0%	0%	0%	0%
5	50.0%	0%	50.0%	0%	0%	0%	0%	0%
6	0%	100%	0%	0%	0%	0%	0%	0%
7	63.4%	23.2%	8.4%	2.4%	1.7%	0.5%	0.2%	0.2%

Table S6

Percentages with symptoms at age 14 sorted by symptoms at age 12

	Age 14: number of symptoms							
Age 12: number of symptoms	0	1	2	3	4	5	6	7
0	74.3%	18.1%	4.3%	2.6%	0.1%	0.5%	0.0%	0.1%
1	46.1%	26.8%	12.3%	7.1%	1.8%	1.9%	2.6%	1.3%
2	26.5%	27.0%	20.8%	12.0%	4.8%	6.9%	0.0%	1.9%
3	14.1%	26.1%	15.0%	6.6%	38.1%	0.0%	0.0%	0.0%
4	0%	0%	40.7%	34.3%	0%	25.0%	0%	0%
5	0%	0%	0%	0%	44.1%	0%	55.9%	0%
6	0%	0%	0%	0%	100%	0%	0%	0%
7	61.1%	20.6%	8.3%	4.9%	2.0%	2.0%	0.8%	0.5%

Supplementary Tables Study II (S7-S25)

Table S7

Pearson Product Moment Correlations between Depression and Neuroticism, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Depression – age 10	1								
2. Depression – age 12	.35***	1							
3. Depression – age 14	.38***	.56***	1						
4. Depression – age 16	.21**	.22*	.32***	1					
5. Neuroticism – age 10	.27***	.21***	.21***	.03	1				
6. Neuroticism – age 12	.32***	.37***	.33***	.12*	.45***	1			
7. Neuroticism – age 14	.37***	.38***	.53***	.20***	.36***	.55***	1		
8. Neuroticism – age 16	.26***	.30***	.33***	.36***	.29***	.42***	.55***	1	
9. Gender[a]	.03	.01	.14**	.04	-.01	-.05	.16**	.26***	1

Note. *indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S8

Pearson Product Moment Correlations between Depression and Extraversion, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Depression – age 10	1								
2. Depression – age 12	.35***	1							
3. Depression – age 14	.38***	.56***	1						
4. Depression – age 16	.21**	.22*	.32***	1					
5. Extraversion – age 10	-.15***	-.07	-.09*	-.03	1				
6. Extraversion – age 12	-.19***	-.20***	-.18***	-.16**	.48***	1			
7. Extraversion – age 14	-.23***	-.21***	-.28***	-.17**	.40***	.59***	1		
8. Extraversion – age 16	-.15***	-.16***	-.19***	-.21***	.38***	.50***	.62***	1	
9. Gender[a]	.03	.01	.14**	.04	.01	.12*	.02	.01	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S9

Pearson Product Moment Correlations between Depression and Conscientiousness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Depression – age 10	1								
2. Depression – age 12	.35***	1							
3. Depression – age 14	.38***	.56***	1						
4. Depression – age 16	.21**	.22*	.32***	1					
5. Conscientiousness – age 10	-.21***	-.10*	-.14**	-.07	1				
6. Conscientiousness – age 12	-.20***	-.23***	-.22***	-.20***	.49***	1			
7. Conscientiousness – age 14	-.22***	-.26***	-.34***	-.19***	.41***	.65***	1		
8. Conscientiousness – age 16	-.12*	-.17**	-.16***	-.23***	.30***	.50***	.64***	1	
9. Gender[a]	.03	.01	.14**	.04	.02	.03	.07	.11*	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S10

Pearson Product Moment Correlations between Depression and Agreeableness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Depression – age 10	1								
2. Depression – age 12	.35***	1							
3. Depression – age 14	.38***	.56***	1						
4. Depression – age 16	.21**	.22*	.32***	1					
5. Agreeableness – age 10	-.11*	-.06	-.04	-.05	1				
6. Agreeableness – age 12	-.20**	-.16***	-.21***	-.13*	.37***	1			
7. Agreeableness – age 14	-.14*	-.06	-.18***	-.15**	.36***	.56***	1		
8. Agreeableness – age 16	-.12*	-.06	-.10	-.15**	.28***	.43***	.58***	1	
9. Gender[a]	.03	.01	.14**	.04	.03	.03	.04	.14**	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S11

Pearson Product Moment Correlations between Depression and Openness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Depression – age 10	1								
2. Depression – age 12	.35***	1							
3. Depression – age 14	.38***	.56***	1						
4. Depression – age 16	.21**	.22*	.32***	1					
5. Openness – age 10	.05	.07	.07	.06	1				
6. Openness – age 12	.00	-.07	-.06	.06	.51***	1			
7. Openness – age 14	.01	.01	-.07	.09	.40***	.58***	1		
8. Openness – age 16	.09	.10	.10	.14*	.31***	.38***	.61***	1	
9. Gender[a]	.03	.01	.14**	.04	.06	.16**	-.03	-.15**	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S12

Pearson Product Moment Correlations between Neuroticism and Extraversion, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Neuroticism – age 10	1								
2. Neuroticism – age 12	.45***	1							
3. Neuroticism – age 14	.36***	.55***	1						
4. Neuroticism – age 16	.29***	.42***	.55***	1					
5. Extraversion – age 10	-.34***	-.16***	-.11**	-.12**	1				
6. Extraversion – age 12	-.24***	-.38***	-.20***	-.10*	.48***	1			
7. Extraversion – age 14	-.19***	-.23***	-.36***	-.17***	.40***	.59***	1		
8. Extraversion – age 16	-.16***	-.21***	-.19***	-.32***	.38***	.50***	.62***	1	
9. Gender[a]	-.01	-.05	.16**	.26***	.01	.12*	.02	.01	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S13

Pearson Product Moment Correlations between Neuroticism and Conscientiousness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Neuroticism – age 10	1								
2. Neuroticism – age 12	.45***	1							
3. Neuroticism – age 14	.36***	.55***	1						
4. Neuroticism – age 16	.29***	.42***	.55***	1					
5. Conscientiousness – age 10	-.39***	-.23***	-.24***	-.10*	1				
6. Conscientiousness – age 12	-.31***	-.42***	-.28***	-.16*	.49***	1			
7. Conscientiousness – age 14	-.20***	-.29***	-.39***	-.18***	.41***	.65***	1		
8. Conscientiousness – age 16	-.17***	-.20***	-.21***	-.24***	.30***	.50***	.64***	1	
9. Gender[a]	-.01	-.05	.16**	.26***	.04	.10	-.02	.09	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S14

Pearson Product Moment Correlations between Neuroticism and Agreeableness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Neuroticism – age 10	1								
2. Neuroticism – age 12	.45***	1							
3. Neuroticism – age 14	.36***	.55***	1						
4. Neuroticism – age 16	.29***	.42***	.55***	1					
5. Agreeableness – age 10	-.36***	-.28***	-.16***	-.17***	1				
6. Agreeableness – age 12	-.16***	-.45***	-.26***	-.22***	.37***	1			
7. Agreeableness – age 14	-.11*	-.25***	-.33***	-.22***	.36***	.56***	1		
8. Agreeableness – age 16	-.03	-.12*	-.15**	-.27***	.28***	.43***	.58***	1	
9. Gender[a]	-.01	-.05	.16**	.26***	.09	.06	.02	.10*	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S15

Pearson Product Moment Correlations between Neuroticism and Openness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Neuroticism – age 10	1								
2. Neuroticism – age 12	.45***	1							
3. Neuroticism – age 14	.36***	.55***	1						
4. Neuroticism – age 16	.29***	.42***	.55***	1					
5. Openness – age 10	-.14***	-.07	.01	.09*	1				
6. Openness – age 12	-.10*	-.17***	-.04	.06	.51***	1			
7. Openness – age 14	-.05	-.10*	-.03	.06	.40***	.58***	1		
8. Openness – age 16	-.05	-.06	.01	-.00	.31***	.38***	.61***	1	
9. Gender[a]	-.01	-.05	.16**	.26***	.06	.16**	-.03	-.15**	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S16

Pearson Product Moment Correlations between Extraversion and Conscientiousness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Extraversion – age 10	1								
2. Extraversion – age 12	.48***	1							
3. Extraversion – age 14	.40***	.59***	1						
4. Extraversion – age 16	.38***	.50***	.62***	1					
5. Conscientiousness – age 10	.32***	.22***	.26***	.17***	1				
6. Conscientiousness – age 12	.17***	.39***	.29***	.17***	.49***	1			
7. Conscientiousness – age 14	.14***	.24***	.32***	.18***	.41***	.65***	1		
8. Conscientiousness – age 16	.13**	.15***	.17***	.20***	.30***	.50***	.64***	1	
9. Gender[a]	.01	.12*	.02	.01	.04	.10	-.02	.09	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S17

Pearson Product Moment Correlations between Extraversion and Agreeableness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Extraversion – age 10	1								
2. Extraversion – age 12	.48***	1							
3. Extraversion – age 14	.40***	.59***	1						
4. Extraversion – age 16	.38***	.50***	.62***	1					
5. Agreeableness – age 10	.32***	.21***	.24***	.17***	1				
6. Agreeableness – age 12	.20***	.41***	.29***	.20***	.37***	1			
7. Agreeableness – age 14	.14***	.25***	.36***	.20***	.36***	.56***	1		
8. Agreeableness – age 16	.07	.19***	.20***	.28***	.28***	.43***	.58***	1	
9. Gender[a]	.01	.12*	.02	.01	.09	.06	.02	.10*	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S18

Pearson Product Moment Correlations between Extraversion and Openness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Extraversion – age 10	1								
2. Extraversion – age 12	.48***	1							
3. Extraversion – age 14	.40***	.59***	1						
4. Extraversion – age 16	.38***	.50***	.62***	1					
5. Openness – age 10	.37***	.20***	.13**	.12**	1				
6. Openness – age 12	.25***	.39***	.21***	.13**	.51***	1			
7. Openness – age 14	.16***	.20***	.24***	.13***	.40***	.58***	1		
8. Openness – age 16	.13**	.07	.12**	.18***	.31***	.38***	.61***	1	
9. Gender[a]	.01	.12*	.02	.01	.06	.16**	-.03	-.15**	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S19

Pearson Product Moment Correlations between Conscientiousness and Agreeableness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Conscientiousness – age 10	1								
2. Conscientiousness – age 12	.49***	1							
3. Conscientiousness – age 14	.41***	.65***	1						
4. Conscientiousness – age 16	.30***	.50***	.64***	1					
5. Agreeableness – age 10	.52***	.26***	.28***	.23***	1				
6. Agreeableness – age 12	.29***	.51***	.38***	.29***	.37***	1			
7. Agreeableness – age 14	.22***	.31***	.45***	.30***	.36***	.56***	1		
8. Agreeableness – age 16	.19***	.22***	.26***	.38***	.28***	.43***	.58***	1	
9. Gender[a]	.04	.10	-.02	.09	.09	.06	.02	.10*	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S20

Pearson Product Moment Correlations between Conscientiousness and Openness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Conscientiousness – age 10	1								
2. Conscientiousness – age 12	.49***	1							
3. Conscientiousness – age 14	.41***	.65***	1						
4. Conscientiousness – age 16	.30***	.50***	.64***	1					
5. Openness – age 10	.27***	.06	.00	-.04	1				
6. Openness – age 12	.15***	.23***	.09*	.00	.51***	1			
7. Openness – age 14	.10*	.10*	.15***	-.03	.40***	.58***	1		
8. Openness – age 16	.03	-.04	-.01	-.03	.31***	.38***	.61***	1	
9. Gender[a]	.04	.10	-.02	.09	.06	.16**	-.03	-.15**	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S21

Pearson Product Moment Correlations between Agreeableness and Openness, ages 10-16

	1	2	3	4	5	6	7	8	9
1. Agreeableness – age 10	1								
2. Agreeableness – age 12	.37***	1							
3. Agreeableness – age 14	.36***	.56***	1						
4. Agreeableness – age 16	.28***	.43***	.58***	1					
5. Openness – age 10	.31***	.06	.04	.00	1				
6. Openness – age 12	.17***	.27***	.11*	-.06	.51***	1			
7. Openness – age 14	.07	.10*	.10*	-.05	.40***	.58***	1		
8. Openness – age 16	.02	-.04	-.05	-.03	.31***	.38***	.61***	1	
9. Gender[a]	.09	.06	.02	.10*	.06	.16**	-.03	-.15**	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$, [a]0 = male; 1 = female.

Table S22

Random Intercept Cross-lagged Panel Model Analyses of Depressive Symptoms and each of the Big Five Personality Traits, ages 10-16

Standardized slope coefficients (*p*-value) B [95% CI]					
Personality traits → Depression					
	Neuroticism → Depression	Extraversion → Depression	Conscientiousness → Depression	Agreeableness → Depression	Openness → Depression
Ages 10-12	**.15 (*p*=.002) [.06, .24]**	-.03 (*p*=.372) [-.09, .06]	-.05 (*p*=.256) [-.13, .03]	**-.08 (*p*=.024) [-.16, -.01]**	.03 (*p*=.498) [-.05, .10]
Ages 12-14	**.13 (*p*=.002) [.05, .21]**	-.03 (*p*=.372) [-.08, .03]	-.04 (*p*=.256) [-.11, .03]	**-.07 (*p*=.024) [-.12, -.01]**	.02 (*p*=.498) [-.04, .08]
Ages 14-16	.03 (*p*=.685) [-.11, .17]	-.05 (*p*=.372) [-.15, .06]	-.07 (*p*=.256) [-.19, .05]	**-.11 (*p*=.024) [-.20, -.02]**	.04 (*p*=.498) [-.07, .14]
Depression → Personality Traits					
	Depression → Neuroticism	Depression → Extraversion	Depression → Conscientiousness	Depression → Agreeableness	Depression → Openness
Ages 10-12	**.23 (*p*<.001) [.13, .34]**	**-.07 (*p*=.027) [-.13, -.01]**	-.10 (*p*=.161) [-.24, .03]	**-.17 (*p*=.022) [-.33, -.02]**	.02 (*p*=.605) [-.04, .07]
Ages 12-14	**.26 (*p*<.001) [.15, .34]**	**-.08 (*p*=.027) [-.15, -.01]**	**-.15 (*p*=.009) [-.27, -.04]**	.02 (*p*=.709) [-.08, .17]	.02 (*p*=.605) [-.05, .09]
Ages 14-16	.08 (*p*=.245) [-.06, .22]	**-.08 (*p*=.027) [-.16, -.01]**	.07 (*p*=.156) [-.02, .15]	-.01 (*p*=.911) [-.12, .12]	.02 (*p*=.605) [-.07, .11]

Note. All lags between a specific personality trait and depression were set to be equal, with the following exceptions: Neuroticism predicting depression (predisposition), and depression predicting neuroticism (scar) were set to be free from ages 14 to 16, and depression predicting conscientiousness (scar) and agreeableness (scar) were set to be freely estimated across all ages.

Table S23

Model Comparison of Random Intercept Cross-lagged Panel Models for Depressive Symptoms and each of the Big Five Personality Traits, Neuroticism and Gender not accounted for. Testing Whether Predictions are Similar or Different across ages 10-16

	χ^2	df	CFI	RMSEA	90% CI RMSEA	Δdf	$\Delta\chi^2$ (p-value)
Neuroticism → depression and depression → neuroticism							
All cross-lagged free	45.11	9	.955	.070	.051-.091		
All cross-lagged fixed vs. all cross-lagged free	59.05	13	.943	.066	.049-.083	4	14.47 (.006)
All cross-lagged fixed except d/n14 to d/n16 vs. all free	**44.68**	**11**	**.958**	**.061**	**.043-.080**	**2**	**1.04 (.594)**
Extraversion → depression and depression → extraversion							
All cross-lagged free	33.38	9	.967	.058	.038-.079		
All cross-lagged fixed vs. all free	**37.92**	**13**	**.966**	**.048**	**.031-.067**	**4**	**4.28 (.370)**
Conscientiousness → depression and depression → conscientiousness							
All cross-lagged free	31.92	9	.971	.056	.036-.077		
All cross-lagged fixed vs. all free	50.54	13	.925	.059	.043-.077	4	14.02 (.007)
All cross-lagged fixed except d10/12/14 to c12/14/16 vs. all free	**33.94**	**11**	**.971**	**.051**	**.032-.070**	**2**	**1.84 (.399)**
Agreeableness → depression and depression → agreeableness							
All cross-lagged free	23.42	9	.975	.044	.023-.067		
All cross-lagged fixed vs. all free	34.90	13	.962	.045	.028-.064	4	11.35 (.023)
All cross-lagged fixed except d10/12/14 to a12/14/16 vs. all free	**28.42**	**11**	**.970**	**.044**	**.024-.064**	**2**	**5.04 (.080)**
Openness → depression and depression → openness							
All cross-lagged free	20.37	9	.983	.039	.016-.062		
All cross-lagged fixed vs. all free	**28.10**	**13**	**.978**	**.038**	**.018-.057**	**4**	**7.34 (.120)**

Note. Bold indicates the best fitting model for each personality trait (i.e., when the fixed model did not deteriorate the model fit of the free model, we would keep the fixed model/equal effects across ages). d=depression, n=neuroticism, c=conscientiousness, a=agreeableness.

Table S24

Cross-Lagged Panel Model Analyses of Depressive Symptoms and the Big Five Personality Traits, ages 10-16

Standardized coefficients (*p*-value) B [95% CI]					
Personality traits → Depression					
	Neuroticism → Depression	Extraversion → Depression	Conscientiousness → Depression	Agreeableness → Depression	Openness → Depression
Ages 10-12	**.15 (*p*=.002) [.06, .24]**	-.05 (*p*=.055) [-.09, .00]	**-.05 (*p*=.031) [-.10, -.01]**	**-.07 (*p*=.026) [-.13, -.01]**	**.05 (*p*=.021) [.01, .10]**
Ages 12-14	**.12 (*p*=.002) [.04, .21]**	-.04 (*p*=.055) [-.08, .00]	**-.04 (*p*=.031) [-.08, -.00]**	**-.05 (*p*=.026) [-.10, -.01]**	**.04 (*p*=.021) [.01, .08]**
Ages 14-16	.02 (*p*=.810) [-.13, .16]	-.07 (*p*=.055) [-.13, -.00]	**-.07 (*p*=.031) [-.14, -.01]**	**-.08 (*p*=.026) [-.16, -.01]**	**.07 (*p*=.021) [.01, .12]**
Depression → Personality Traits					
	Depression → Neuroticism	Depression → Extraversion	Depression → Conscientiousness	Depression → Agreeableness	Depression → Openness
Ages 10-12	**.21 (*p*<.001) [.11, .31]**	**-.09 (*p*<.001) [-.13, -.05]**	**-.11 (*p*<.001) [-.16, -.06]**	**-.14 (*p*=.004) [-.24, -.03]**	.04 (*p*=.112) [-.01, .09]
Ages 12-14	**.22 (*p*<.001) [.13, .32]**	**-.11 (*p*<.001) [-.15, -.05]**	**-.12 (*p*<.001) [-.18, -.06]**	.02 (*p*=.497) [-.04, .09]	.04 (*p*=.112) [-.01, .10]
Ages 14-16	.07 (*p*=.290) [-.06, .20]	**-.11 (*p*<.001) [-.16, -.06]**	.05 (*p*=.160) [-.02, .13]	.03 (*p*=.497) [-.05, .10]	.05 (*p*=.112) [-.01, .12]

Note. All lags between a specific personality trait and depression were set to be equal, with the following exceptions: Neuroticism predicting depression (predisposition), and depression predicting neuroticism (scar) were set to be free from ages 14 to 16, depression predicting conscientiousness (scar) were set to be free from ages 14 to 16, and depression predicting agreeableness (scar) were set to be free from ages 10 to 12.

Table S25

Model Comparison of Cross-lagged Panel Models for Depressive Symptoms and each of the Big Five Personality Traits. Testing Whether Predictions are Similar or Different across ages 10-16

	χ^2	df	CFI	RMSEA	90% CI RMSEA	Δdf	$\Delta\chi^2$ (p-value)
Neuroticism → depression and depression → neuroticism							
All cross-lagged free	57.22	12	.950	.068	.051, .086		
All cross-lagged fixed vs. all cross-lagged free	78.03	16	.931	.069	.054, .085	4	20.88 (<.001)
All cross-lagged fixed except d/n14 to d/n16 vs. all free	**58.66**	**14**	**.951**	**.063**	**.047, .080**	**2**	**1.16 (.559)**
Extraversion → depression and depression → extraversion							
All cross-lagged free	110.78	29	.954	.059	.048, .071		
All cross-lagged fixed vs. all free	**116.51**	**33**	**.954**	**.056**	**.045, .067**	**4**	**6.19 (.186)**
Conscientiousness → depression and depression → conscientiousness							
All cross-lagged free	88.81	29	.967	.050	.039, .063		
All cross-lagged fixed vs. all free	104.31	33	.960	.052	.041, .063	4	15.80 (.003)
All cross-lagged fixed except d14 to c16 vs. all free	**92.06**	**32**	**.966**	**.048**	**.037, .060**	**3**	**2.82 (0.420)**
Agreeableness → depression and depression → agreeableness							
All cross-lagged free	101.68	29	.955	.056	.044, .067		
All cross-lagged fixed vs. all free	113.88	33	.950	.055	.044, .066	4	12.53 (.014)
All cross-lagged fixed except a5 to d6 and d4 to a5 vs. all free	**103.08**	**32**	**.956**	**.052**	**.041, .064**	**3**	**3.30(0.348)**
Openness → depression and depression → openness							
All cross-lagged free	86.81	29	.963	.050	.038, .062		
All cross-lagged fixed vs. all free	**91.92**	**33**	**.963**	**.047**	**.036, .058**	**4**	**5.03 (.285)**

Note. Bold indicates the best fitting model for each personality trait (i.e., when the fixed model did not deteriorate the model fit of the free model, we would keep the fixed model/equal effects across ages). d=depression, n=neuroticism, c=conscientiousness, a=agreeableness.

Supplementary materials Study III (List of Stressful Life Events and Table S26)

List of Stressful Life Events

1. *Have one or more children that are not your siblings come to live in your home?*
2. *Have you gotten new half- or full siblings?*
3. *Have your parents separated or divorced and no longer live together?*
4. *Did a new adult move into your home, e.g., a new girl/boy-friend or partner to your parents?*
5. *Did such a new adult move out of your home, e.g., a new girl/boy-friend or partner?*
6. *Have you moved to a new place?*
7. *Have you changed your school?*
8. *Have your parents been evaluated at or admitted to a hospital for more than 24 hours?*
9. *Have either of your parents suffered a life-threatening illness which made them unable to work?*
10. *Have either of your parents been admitted to a hospital for more than 7 days?*
11. *Have you lost contact with someone important to you (such as a friend or relative)?*
12. *Have you and a best friend parted because you had a quarrel or fight (seems to be permanent)?*
13. *Have you and a girl/boy-friend broken up because you fought, were no longer in love, or because one of you moved?*
14. *Did a pet you felt close to die?*
15. *Have either of your parents been in prison?*
16. *Has your family's income been substantially less than usual?*
17. *Have you been forced to live elsewhere than your home for a period?*
18. *Have you ever gotten very sick (risk of death or a serious illness)?*
19. *Have you ever been admitted to a hospital for more than 24 hours?*
20. *Have you ever been in a serious accident? (fire/burn, poisoning, traffic-accident, near drowning, bitten by an animal, ect.?)*
21. *Has anyone close to you died (parents or full- half- and step- siblings)?*
22. *Have you ever experienced a serious natural disaster, such as a storm, hurricane, flood, earthquake?*
23. *Have you ever been in a terrible fire?*
24. *Have you ever seen or heard something horrible happen to anyone?*
25. *Have you ever heard about something horrible happen to someone you care about, but not seen it yourself?*
26. *Have you ever had contact with something that could make you very sick or die?*
27. *Have you or someone you were with ever hurt another person seriously?*
28. *Have you ever been hit or hurt by somebody?*
29. *Have anyone in your family ever hit or hurt you badly?*
30. *Have you ever been kidnapped or locked up against your will?*
31. *Have you ever been exposed to sexual abuse?*

Note. Parents and children were asked to affirm if the following events occurred during the last two years.

Table S26

Pearson Product Moment Correlation between Depression, Stressful Life Events, and Bullying Victimization, ages 8-14

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Gender[a]	1												
2. Depression – age 8	-.11*	1											
3. Depression – age 10	.02	.31***	1										
4. Depression – age 12	.09*	.37***	.38***	1									
5. Depression – age 14	.14**	.26***	.36***	.55***	1								
6. Stressful life events – age 8	-.05	.05	-.01	.08	.10	1							
7. Stressful life events – age 10	.02	-.00	.08	.01	.12**	.38***	1						
8. Stressful life events – age 12	-.06	.02	.18**	.12**	.23***	.29***	.39***	1					
9. Stressful life events – age 14	.07	.06	.11*	.13*	.22***	.18**	.18**	.45***	1				
10. Bullying victimization – age 8	-.00	.21***	.04	.14*	.18**	.11	.08	.11*	.12	1			
11. Bullying victimization – age 10	-.05*	.19***	.09	.12*	.15**	-.02	-.03	.03	.09	.39***	1		
12. Bullying victimization – age 12	-.11*	.09	.08	.21***	.25***	.02	.03	-.02	.12*	.22***	.33***	1	
13. Bullying victimization – age 14	.03	.11	.07	.12*	.20**	.02	.02	.01	.12*	.21***	.23**	.39***	1

Note. * indicates $p < .05$, ** indicates $p < .01$, *** indicates $p < .001$. [a]0 = male; 1 = female.

ISBN 978-82-326-7630-9 (printed ver.)
ISBN 978-82-326-7629-3 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology