Gender Differences in Depression in Representative National Samples: Meta-Analyses of Diagnoses and Symptoms

Rachel H. Salk
University of Wisconsin – Madison and
University of Pittsburgh School of Medicine

Janet S. Hyde and Lyn Y. Abramson
University of Wisconsin – Madison

In 2 meta-analyses on gender differences in depression in nationally representative samples, we advance previous work by including studies of depression diagnoses and symptoms to (a) estimate the magnitude of the gender difference in depression across a wide array of nations and ages; (b) use a developmental perspective to elucidate patterns of gender differences across the life span; and (c) incorporate additional theory-driven moderators (e.g., gender equity). For major depression diagnoses and depression symptoms, respectively, we meta-analyzed data from 65 and 95 articles and their corresponding national data sets, representing data from 1,716,195 and 1,922,064 people in over 90 different nations. Overall, odds ratio $(OR) = 1.95$, 95% confidence interval (CI) [1.88, 2.03], and $d = 0.27$ [0.26, 0.29]. Age was the strongest predictor of effect size. The gender difference for diagnoses emerged earlier than previously thought, with $OR = 2.37$ at age 12. For both meta-analyses, the gender difference peaked in adolescence $(OR = 3.02$ for ages 13–15, and $d = 0.47$ for age 16) but then declined and remained stable in adulthood. Cross-national analyses indicated that larger gender differences were found in nations with greater gender equity, for major depression, but not depression symptoms. The gender difference in depression represents a health disparity, especially in adolescence, yet the magnitude of the difference indicates that depression in men should not be overlooked.

Keywords: depression, gender, meta-analysis, development, gender equity

Supplemental materials: http://dx.doi.org/10.1037/bul0000102.supp

Depression is a global health priority. According to the World Health Organization (2016), depression accounts for fully 10% of the total nonfatal disease burden worldwide. Moreover, this burden falls disproportionately on girls and women. Recent estimates of the global 12-month prevalence of major depressive disorder were 5.8% in women and 3.5% in men (Ferrari et al., 2013). The gender difference in depression—generally believed to be twice as many women experiencing major depression as men—represents a major health disparity. However, despite assertions that the gender difference in depression is among the most robust of findings in psychopathology research (e.g., Bebbington, 1996), and extensive empirical and theoretical work on gender differences in depression, this large body of sometimes inconsistent research has yet to be synthesized meta-analytically. The current set of meta-analyses advance previous work by including studies of depression diagnoses as well as symptoms to (a) estimate the magnitude of the gender difference in depression; (b) use a developmental lens to elucidate the patterns of gender differences across the life span; and (c) examine theory-driven, conceptually relevant moderators (e.g., nation-level gender equity).

Background

In the 1970s, Myrna Weissman first underscored the gender difference in depression, noting that approximately twice as many women experience depression as men among adults in clinical and community samples (Weissman & Klerman, 1977). Following this landmark article, there was a proliferation of research and theories on gender differences in depression (Bebbington, 1996; Kuehner, 2003; Nolen-Hoeksema, 1987; Piccinelli & Wilkinson, 2000; Weissman & Klerman, 1977; for an overview of explanatory models, see Hammarström, Lehti, Danielsson, Bengs, & Johansson, 2009). In the vast majority of epidemiological reports among

adults, women have higher rates of major depression compared with men; on average, the ratio is 2:1 (Andrade et al., 2003; Bromet et al., 2011). However, findings also suggest that the 2:1 ratio is not universal and may vary substantially across nations. For example, in 18 countries from the WHO World Mental Health Surveys (Kessler & Üstün, 2008), odds ratios (ORs, female/male) for 12-month major depressive episode (MDE[1]) ranged from 1.2 to 2.7 across 18 countries and 89,037 participants (Bromet et al., 2011). Given this variability, it is critical to use meta-analysis to estimate the overall magnitude and consistency of the gender difference in depression across different nations and with different assessments of major depression. Other widely held beliefs about gender differences, such as the gender difference in math performance, have sometimes been found to be inaccurate when the data are meta-analyzed (Hyde, Lindberg, Linn, Ellis, & Williams, 2008a; Lindberg, Hyde, Petersen, & Linn, 2010). Moreover, given evidence of cross-national variations, it is important to understand nation-level variables (e.g., economic development, gender equity) that may account for variability in the magnitude of the gender difference.

In addition to examining variations in the gender difference in depression across nations, it is also critical to take a developmental perspective. Several studies indicate that, among the general population, there is no gender difference or even a somewhat higher prevalence of depression among boys than girls in childhood (Avenevoli, Knight, Kessler, & Merikangas, 2008; Twenge & Nolen-Hoeksema, 2002). The female preponderance in depression is thought to emerge by ages 13–15 (e.g., Hankin et al., 1998; Twenge & Nolen-Hoeksema, 2002; Wade, Cairney, & Pevalin, 2002; Wichstrøm, 1999). However, research on the time course of the emergence of the gender difference in adolescence has been accepted as a fundamental fact in the depression literature when it is actually based on only a few studies. For example, in a landmark article, Hankin and colleagues (1998) found that the gender difference in clinical depression emerged by ages 13–15 and then widened between ages 15 and 18. This conclusion has been widely accepted (the article had been cited 1693 times as of December 2016) based on findings from one sample from one region of New Zealand (see Kessler, McGonagle, Swartz, Blazer, & Nelson, 1993, for the other widely cited study on gender differences in adolescence, based on U.S. data). A meta-analysis on gender differences in depression with a developmental focus is the next major step to pinpoint the time course of the emerging gender difference in depression.

Additionally, developmental patterns of the gender difference beyond adolescence have been largely neglected empirically. The limited findings in adulthood are inconsistent with respect to both the magnitude and direction of the gender difference in depression (Angst et al., 2002; Bebbington et al., 1998; Mirowsky, 1996; Oksuzyan et al., 2010; Patten et al., 2016). Estimates of the gender difference in depression in older adults also suggest marked variability. A meta-analysis of 24 studies among individuals ages 75 and older reported gender ratios between 1.4 and 2.2 (Luppa et al., 2012). It was one of the goals of the current meta-analyses to bring clarity to developmental patterns throughout the life span.

Lastly, despite much attention to the 2:1 ratio for the gender difference in *major depression*, the magnitude of the gender difference in levels of *depression symptoms* in the general population has received less attention. Psychiatric research in the past several decades has focused on the use of diagnostic categories as specified in the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*; American Psychiatric Association, 1980, 1987, 1994, 2000, 2013) and *International Classification of Disease* (*ICD*; World Health Organization, 1992). However, there is increased recognition of the validity (e.g., Markon, Chmielewski, & Miller, 2011) and value (Cuthbert & Insel, 2013) of dimensional assessments, as well as the impairment associated with subthreshold levels of symptoms that do not meet diagnostic criteria. Indeed, adolescents and adults with subthreshold depression symptoms and minor depression still experience significant impairment (e.g., Lewinsohn, Solomon, Seeley, & Zeiss, 2000) and are at elevated risk for later development of major depression and suicidal behaviors (Cuijpers, de Graaf, & van Dorsselaer, 2004; Fergusson, Horwood, Ridder, & Beautrais, 2005; Klein, Shankman, Lewinsohn, & Seeley, 2009). These important subthreshold levels are captured in symptom questionnaires; however, extant research on the magnitude of gender differences in depression symptoms is limited. One meta-analysis reported effect sizes ranging from $d = -0.06$ at age 12 to $d = 0.22$ at ages 14 and 15 (positive values indicate more depression symptoms among girls; Twenge & Nolen-Hoeksema, 2002). Although this study represented a step forward in the research literature, it was limited to the Children's Depression Inventory (CDI; Kovacs, 1985) scale and samples of 8- to 16-year-old participants in the United States and Canada ($n = 43,916$). Given the impairment associated with high levels of depression symptoms in the absence of a diagnosis, it is critical to estimate the magnitude of the gender difference in depression symptoms more comprehensively: throughout the life span, across nations, and with multiple symptom measurements.

Thus, in the current set of meta-analyses using nationally representative samples, we estimated the magnitude of the gender difference in (a) major depression diagnoses and (b) levels of depression symptoms. Moreover, meta-analysis allowed us to chart the developmental course of the gender difference from childhood through late adulthood. Meta-analysis also enabled us to test whether the gender difference is universal across nations or whether there are cultural factors such as gender inequality that account for cross-national variations.

Gender Differences in Depression Across the Life Span

Based on both developmental psychopathology theory and past research (reviewed above), age was used as a moderator in the current set of meta-analyses to answer one of our fundamental questions: What is the pattern of gender differences in depression across the life span?

There now is consensus that the gender difference in depression has a multifactorial etiology (Cyranowski, Frank, Young, & Shear, 2000; Hyde, Mezulis, & Abramson, 2008b). Theories of developmental psychopathology contend that there are multiple pathways

---

[1] Major depressive episode (MDE) and major depressive disorder (MDD) are distinct. MDD requires the presence of a major depressive episode (MDE) and the absence of a manic, mixed, or hypomanic episode. Thus, MDE includes depressive episodes that occur in both unipolar depression and bipolar disorder, whereas MDD includes only unipolar depression. However, the vast majority of lifetime and 12-month MDE is MDD.

to the gender difference in depression involving combinations and interactions of risk factors that span multiple levels of analysis (Cicchetti & Rogosch, 2002). Notably, these pathways to the gender difference in depression occur in a developmental context. Theories highlight how specific vulnerability factors come online at critical developmental periods in adolescence and/or interact with stressors in adolescence to produce the gender difference in adolescence (reviewed by Hyde et al., 2008b). For example, there is a confluence of hormonal and neurodevelopmental changes that vary by sex during the pubertal transition and may influence the gender difference in depression. Thus, a developmental approach is key to understanding patterns in the gender difference across the adolescent transition and to understand if the gender difference persists across the life span.

In contrast to the focus on the emergence of the gender difference in adolescence, researchers have largely ignored development in adulthood when theorizing about and examining gender differences in depression. The field of developmental psychopathology encourages a life span perspective, as the process of adaptation continues from childhood through adulthood (Cicchetti & Rogosch, 2002). For example, with regard to depression, little is known about levels of stress for women compared with men across adulthood, nor about the importance of various life transitions in adulthood. Theorizing about gender differences in depression will be enriched by an understanding of developmental patterns across adulthood and it was one of the goals of these meta-analyses to elucidate those developmental patterns.

Gender Differences in Depression Across Nations

Past research indicates variability in the magnitude of the gender difference in depression across nations. We used sociological theory and social-structural theory to guide our use of nation-level economic and gender equity indicators as moderators in the current set of meta-analyses.

Sociological Theories

Sociological approaches to mental health emphasize the role of poverty, violence, and gender inequality as factors contributing to the gender difference in depression. Abundant evidence suggests a relationship between financial hardship and depression in both sexes (Reiss, 2013). Because of the feminization of poverty (Belle, 1990; Belle & Doucet, 2003), and the link between poverty and depression, gender differences might also be linked to income inequality and a national's overall wealth. Similar to financial hardship, victimization is also related to depression in both men and women. To the extent that women report higher rates of violent victimization, this may contribute to the gender difference in depression (Koss et al., 1994). Lastly, gender inequality is linked to discrimination against women, which may contribute to the gender difference (Belle & Doucet, 2003). Thus, in the current meta-analyses we investigated nation-level economic factors and gender-equity indicators as moderators of the gender differences in depression.

Social-Structural Theory

Eagly and Wood's social-structural theory (Eagly & Wood, 1999; Wood & Eagly, 2012) also provides a framework for understanding the relationship between gender inequality and the magnitude of psychological gender differences. According to the theory, a society's division of labor by gender drives all other psychological gender differences. These gender differences result from individuals' adaptations to the particular restrictions on or opportunities for their gender in their society. The theory predicts that *larger* gender differences should be observed in nations with *more* gender inequality. Evidence for this theory exists for several psychological gender differences, including mate preferences, mathematics performance, and some aspects of sexuality (Eagly & Wood, 1999; Else-Quest, Hyde, & Linn, 2010; Petersen & Hyde, 2010; Zentner & Mitura, 2012). However, other studies have found *smaller* gender differences in nations with more gender inequality. This pattern has been found for outcomes such as self-reports of personality traits and attitudes about mathematics (Else-Quest et al., 2010; Wood & Eagly, 2012). In the current meta-analyses, we sought to determine which of these patterns would occur (larger or smaller gender differences in nations with more gender inequality) when the outcome was gender differences in depression.

Cross-National Variations: Research on Economic and Gender Equity Factors

Research on the relationship between nation-level economic factors and gender differences in depression is sparse. A study including 18 countries from the WHO World Mental Health Surveys (Kessler & Üstün, 2008) reported that the relationship between gender and MDE did not differ significantly between high-income and low- to middle-income countries, suggesting that economic development does not explain the varying magnitudes of gender differences in different countries (Bromet et al., 2011). Alternatively, a different measure of nation-level economic development may be more sensitive in detecting a relationship to the gender difference in depression. In the current set of meta-analyses, we used two different measures of economic development (income category and income inequality; defined below) and included a more complete set of nations to examine the relationship between nation-level economic factors and gender differences in depression more comprehensively.

Nation-level gender equity indicators are increasingly being used in psychological research (Else-Quest & Grabe, 2012); however, few studies have investigated the relationship between nation-level gender equity and gender differences in depression. Two large multination studies have reported conflicting results, finding that the gender gap in depression was *smaller* and *larger*, respectively, in low gender-equity countries compared with high gender-equity countries (Hopcroft & Bradley, 2007; Van de Velde, Huijts, Bracke, & Bambra, 2013). This relationship is especially complex given the multiple available measures of gender equity. We selected domain-specific indicators of gender equity (rather than composite indicators) that should, theoretically, be tied to gender differences in depression (e.g., contraceptive prevalence, representing a woman's ability to control her own reproduction).

Additional Factors Influencing the Gender Difference in Depression

In addition to examining age and cross-national variations in national wealth and gender equity as moderators, we also explored

whether the magnitude of the gender difference in depression varied according to ethnicity (in U.S. samples only) and over time, that is, whether it is growing larger or smaller.

Ethnicity in the United States and Intersectionality

The extant literature indicates that the prevalence of major depression in the United States varies both by gender and by ethnicity (e.g., Breslau, Kendler, Su, Gaxiola-Aguilar, & Kessler, 2005). However, few studies have tested whether gender differences in depression vary by ethnicity. The importance of this question is highlighted in intersectionality theory, which emphasizes that all people belong to multiple social categories and that these categories are intertwined (Cole, 2009; Else-Quest & Hyde, 2016a). According to this approach, the category of gender should not be considered in isolation, but should be analyzed as it intersects with other categories such as ethnicity. Empirical evidence for these assertions is abundant; space does not permit a thorough review here (for reviews, see Else-Quest & Hyde, 2016a, 2016b).

The limited research on ethnicity, gender, and depression in the United States does not indicate variation by ethnicity in the gender difference in depression (Barnes, Keyes, & Bates, 2013; Breslau et al., 2005; Oquendo et al., 2001; Siegel, Aneshensel, Taub, Cantwell, & Driscoll, 1998). Nonetheless, other meta-analyses on gender differences for related constructs have found notable variations across U.S. ethnic groups. For example, a meta-analysis of gender differences in self-esteem found a small difference favoring White men over White women, $d = 0.20$, but no gender difference for African American samples, $d = -0.04$ (Kling, Hyde, Showers, & Buswell, 1999). Therefore, it was important to test whether gender differences in depression vary across U.S. ethnic groups. We did not conduct analyses stratified by ethnicity in other nations because ethnic groups are distinct in each country and often are not reported.

Trends Over Time

One recent narrative review concluded that internalizing problems for girls increased from the late 20th century to the 21st century (Bor, Dean, Najman, & Hayatbakhsh, 2014). The findings for boys were mixed as to whether they experienced an increase. In contrast, Seedat and colleagues (2009) found a significant narrowing in the gender difference in depression in recent cohorts. Therefore, we tested meta-analytically whether gender differences in depression are widening or narrowing over time.

Sampling Issues

The current set of meta-analyses synthesized data from representative samples, based on an approach pioneered by Hedges and Nowell (1995). They argued that the strongest scientific evidence about gender differences does not come from small studies of convenience samples, but instead comes from larger studies based on representative samples of populations. The Hedges and Nowell strategy has since been used in other meta-analyses on gender differences in cognitive abilities (e.g., Else-Quest et al., 2010; Reilly, Neumann, & Andrews, 2015). Beginning around 1990, with the formation of cross-national collaboration groups studying psychiatric epidemiology (e.g., Cross National Collaborative Group, 1992), data sets based on representative samples became available for gender differences in depression. We were, therefore, able to use this strong methodology for the current meta-analyses.

The Current Study

Given the abundance of available research on gender differences in major depression and in depression symptoms, a meta-analysis is possible and is needed to address the following key questions:

1. How large is the gender difference in major depression? How large is the gender difference in levels of depression symptoms?

2. Following from developmental psychopathology approaches, what is the pattern of gender differences in depression across the life span? How does the direction or magnitude of the gender difference change across development (i.e., at what ages do gender differences appear or disappear, widen or narrow)?

3. Guided by sociological and social-structural theory, does the magnitude of gender differences vary as a function of the nations' gender equity or wealth?

4. Following from an intersectionality approach, are there variations across U.S. ethnic groups in the direction or magnitude of these gender differences?

5. Have gender differences in depression widened over time, that is, across cohorts from the 1970s to 2014?

Method

Identification of Studies and Data Sets

Database searches. Computerized database searches of PsycINFO and PubMed were used to generate an initial pool of potential articles. To identify all relevant articles and related data sets, the following search terms (selected in consultation with a university librarian) were used in PsycINFO and PubMed, respectively: (depression OR depressive OR depressed) AND (sex OR gender); depression AND (gender OR sex OR sex factors).[2] The search terms were optimized for each database (e.g., using MESH terms in PubMed) and were conceptually similar in terms of article yield. Search limits restricted the results to articles that discussed research with human populations and that were published between 1970 and October 4, 2016 (including online first publications). The year 1970 was chosen as the earliest to capture reasonably contemporary research with modern symptom measures and diagnoses from structured interviews based on the *DSM* and *ICD*. PsycINFO and PubMed identified 29,003 and 28,383 articles, respectively,

---

[2] The broadest search term would only have included search terms related to depression. This would ensure that the identified articles included studies that were not focused on gender differences but still reported the relevant statistics. However, PsycINFO and PubMed each identified over 100,000 articles when the search term was solely "depression," leading us to narrow the search to both depression and (gender or sex) in the search terms.

which were considered for inclusion. In this section and throughout this article, we follow MARS reporting standards (American Psychological Association Publications and Communications Board, 2008; see also Moher, Liberati, Tetzlaff, Altman, & the PRISMA Group, 2009).

Abstract processing. Abstracts and citations were imported into Endnote citation manager. Duplicates were deleted, resulting in 46,512 abstracts (see Figure 1). The abstracts were examined for relevant content. At this stage, we included any studies with potentially relevant depression data and, to ensure the quality of sampling, were based on a nationally representative dataset. Abstracts were excluded for any of the following reasons: (a) the sample was not nationally representative (e.g., clearly a community study or a convenience sample); (b) the sample consisted of only one gender; (c) the study reported no empirical data (e.g., a review article); (d) the research was qualitative; (e) the research was conducted on nonhumans; (f) the participants in the study were younger than seven years old (this age cut-off was selected because, for the sake of uniformity, we included only self-report measures of depression symptoms and not parent or teacher report; we did not restrict the age range in the computerized database searches to avoid missing articles that were not tagged with an age); and (g) the abstract did not mention depression or a related construct (e.g., anxiety, stress, internalizing, emotion, psychological distress, psychiatric disorder, or mental health). There were 44,431 abstracts excluded because of the aforementioned reasons, resulting in 2,081 remaining articles (see Figure 1 for additional information).

Article processing. The pdfs from these 2,081 articles were retrieved and examined to determine whether the articles met the criteria for inclusion. At this stage, we excluded studies that were not based on national probability sampling. In other words, we included only population-based surveys representative of the country. We excluded national samples of college students, employees, veterans, twins, primary care patients, and married couples, as these samples do not represent the general population. We excluded representative samples that were limited to one large city or region or even several regions (if they were not randomly selected). We also excluded samples of inpatients or outpatients as this sampling strategy is vulnerable to the criticism that the study is detecting a gender difference in help seeking rather than an actual gender difference in depression (Nolen-Hoeksema, 1987; Pattyn, Verhaeghe, & Bracke, 2015). Nationally representative samples do include individuals currently receiving mental health treatment (unless they are institutionalized) or individuals with a history of receiving mental health treatment, so those individuals were not excluded.

Also, to ensure quality, studies that did not meet the measurement criteria were excluded at this stage. Studies were excluded if their measurement of depression *symptoms* did not meet the following criteria: (a) minimum of three items; (b) self-report; (c) Cronbach's $\alpha \geq .70$ (if provided); and (d) valid and reliable measure of depression based on previously published research[3] (e.g., we excluded studies that used a general measure of psychological distress or negative affect). If a study used a measure that combined anxiety and depression subscales, we contacted the authors to obtain the data solely for the depression subscale.

Studies were excluded in the processing of articles if their measurement of depression *diagnoses* did not include a diagnostic

interview with the participant. Thus, we excluded studies reporting depression diagnoses from the following sources: health insurance claims databases, participants' self-report of physician-diagnosed depression, antidepressant use, and cut-off scores on depression symptom measures (e.g., a cut-off on the CES-D). We contacted authors who reported diagnoses based on symptom cut-off scores to obtain the continuous symptom data for the depression symptom meta-analysis.

If a particular sample of participants was used in more than one article, which was often the case with these national data sets, to maintain independence of samples, we selected the article that had the most complete data (including information on moderator variables such as age and ethnicity) and/or the largest sample size. For nationally representative *longitudinal* studies with multiple waves of data (e.g., Add Health), we included only one wave of data to maintain independence of samples. In these cases, we selected the article with baseline data (whenever possible) to obtain the largest sample size and avoid bias because of attrition.

Additional searches and author contact. If an article provided insufficient information for effect size calculations, we used three strategies to obtain relevant data for that particular national dataset: (a) we conducted computerized database searches using the dataset name and/or authors; (b) we searched the national data set websites for published tables with depression data; and (c) if the study assessed relevant information (e.g., reported on depression symptoms but did not provide the data separately for men and women), all authors of the study for whom we could find email addresses from the article, the Web directory of the authors' academic institution, or a Google search, were contacted. Given our strong interest in age and ethnicity as moderator variables, we also contacted authors for data on gender differences in depression by age and, for U.S. samples, ethnicity if that information was not provided in the original article. We received relevant information for 103 out of the 186 articles for which we contacted authors.

Final sample of studies. Overall, 112 articles from the original search met criteria for inclusion. We added 46 new articles that were not in the original search from additional searching for nationally representative data sets.

The final sample of studies (see Figure 1) for the meta-analyses included data from 65 (diagnosis meta-analysis) and 95 (symptom meta-analysis) articles and their corresponding data sets. Two articles were used in both meta-analyses (Graham, Massak, Demers, & Rehm, 2007; Maske et al., 2016); several samples were used in both meta-analyses, for example, MIDUS. See Tables 1 and 2 for a list of all studies.

The 65 articles (59 in peer-reviewed journals, 6 online publications from national database websites) and their corresponding data sets for the meta-analysis on depression diagnoses provided

---

[3] Some researchers have questioned the construct validity of self-report depression symptom questionnaires, suggesting that these measures may assess general distress and not specifically depression in the general population (e.g., Kendall, Hollon, Beck, Hammen, & Ingram, 1987). The CES-D, which was the most frequently used symptom questionnaire in the depression symptom meta-analysis, was not designed for clinical diagnoses; however, the items are based on symptoms of major depression. Numerous validation studies are available; in one, the CES-D had a sensitivity of 100% and a specificity of 88% for 1-month major depression diagnoses (Beekman et al., 1997). The CES-D had a weighted sensitivity of 40% for all anxiety disorders in the past year, suggesting specificity for depression versus anxiety.
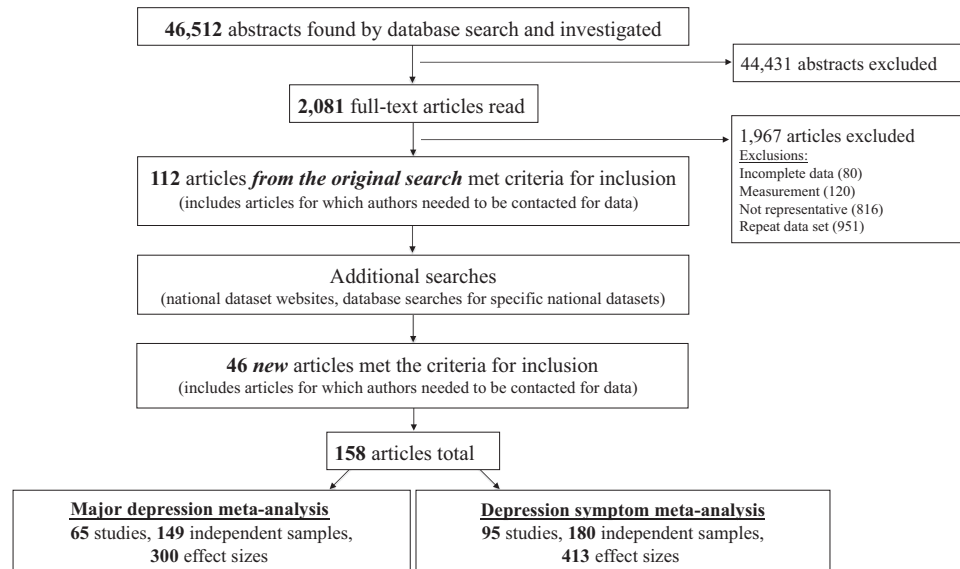
Figure 1. Flowchart of the search and selection procedure.

data comprising 149 samples (this number includes different countries from the same international dataset as well as different years of data collection for cross-sectional national data sets), 300 effect sizes (this number includes separate effect sizes for different age groups and ethnicities), and 1,716,195 people (53% female). These studies were published from 1993 to 2016 with data from 1991–2014 collected in 75 different countries.

The 95 articles (92 in peer-reviewed journals, 2 government publications, and 1 personal communication) and their corresponding data sets for the meta-analysis on depression symptoms utilized data comprising 180 samples, 413 effect sizes, and 1,922,064 people (52% female). These studies were published from 1991 to 2016 and yielded data from 1978–2014 collected from 53 different countries.

Coding of Studies

Studies were coded for information to compute effect sizes and for moderator variables.[4] We double coded 30 studies to compute interrater agreement. The following variables were coded: (a) age of the participants (measured on a continuous scale, using the reported mean or the midpoint of the age range), $r = .99$; (b) country, $\kappa = 1.00$, was used to identify national economic indicators and national gender equity indicators; (c) for U.S. samples, predominant ($\geq 85\%$) ethnic group of participants (White, African American, Asian American, Hispanic, Native American, mixed), $\kappa = 1.00$; and (d) year of data collection, $r = 1.00$.

In any meta-analysis, there is a potential concern that the identified studies have a publication bias, that is, a bias toward publishing studies that found significant and larger gender differences. If present, this bias could mean that the magnitude of the gender difference is overestimated because studies finding no difference are missing from the sample of studies. As one of several ways to address this potential concern, we coded the focus of the article (gender, depression, and other), $\kappa = .84$. If

the majority of the articles were not focused on gender, we would not be concerned about publication bias in regard to gender. We further tested whether the magnitude of the gender difference in depression varied as a function of the article focus to determine if articles that focused on gender reported larger gender differences than articles focused on depression. Gender was identified as the focus of the article if "gender" or a related term was included in the title. If gender was not identified as the focus, depression was identified as the focus of the article if "depress*" or a related term was included in the title, abstract, or keywords. If neither gender nor depression were the focus, "other" was coded.

The following measurement characteristics were also coded: (a) the type of measure used to assess depression diagnoses (diagnostic interview; e.g., CIDI, DIS) or depression symptoms (self-report measure; e.g., CES-D, BDI), $\kappa = 1.0$; (b) the manual used to diagnose depression (*DSM–IV* or *DSM–IV–TR*[5]; *DSM–III* or *DSM–III–R*; or *ICD-10*), $\kappa = 1.0$; (c) type of depression diagnosis (depressive episode vs. depressive disorder), $\kappa = .87$; and (d) the time span for the depression diagnosis (current, 1 month, 6 months, 12 months, and lifetime), $\kappa = 1.00$. If multiple time spans for the depression diagnosis were

---

[4] The majority of studies provided sample sizes for each gender and for each gender stratified by moderating variables; however, for several studies in which this information was not available, we estimated subgroup sample sizes from the total sample size.

[5] All studies in the meta-analysis were conducted with *DSM–IV–TR* or earlier so we rely on it as the source. In *DSM–IV–TR* and *DSM–5* (American Psychiatric Association, 2013), the criteria for a major depressive episode are nearly identical. The one exception is that the bereavement exclusion criterion was removed in *DSM–5*. In the *DSM–IV–TR*, criterion E for a major depressive episode specified "the symptoms are not better accounted for by Bereavement, i.e., after the loss of a loved one, the symptoms persist for longer than 2 months or are characterized by marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation" (American Psychiatric Association, 2000, p. 356).

Table 1

Studies of Gender Differences in Major Depression Diagnoses

Study	OR	Country	NM	NF	Interview	Year	Age	Sample	E	F	M	T
Alaimo et al. (2002)	2.38	U.S.	365	389	DIS	1991	15–16	Third National Health and Nutrition Examination Survey (NHANES III)	1	1	2	4
Andrews et al. (2001)[a]	2.40	Australia	461	608	CIDI 2.1	1997	18–24	Australian National Survey of Mental Health & Well-Being (NSMHWB)		3	1	3
Andrews et al. (2001)[a]	1.97	Australia	940	1,249	CIDI 2.1	1997	25–34	NSMHWB		3	1	3
Andrews et al. (2001)[a]	1.97	Australia	1,131	1,349	CIDI 2.1	1997	35–44	NSMHWB		3	1	3
Andrews et al. (2001)[a]	1.69	Australia	815	1,017	CIDI 2.1	1997	45–54	NSMHWB		3	1	3
Andrews et al. (2001)[a]	2.14	Australia	627	652	CIDI 2.1	1997	55–64	NSMHWB		3	1	3
Andrews et al. (2001)[a]	2.78	Australia	731	1,061	CIDI 2.1	1997	65+	NSMHWB		3	1	3
Arokiasamy et al. (2013)	1.41	India	1,042	3,620	CIDI	2008	18–49	WHO Study on global Aging and adult health (SAGE) - India		3	3[b]	3
Arokiasamy et al. (2013)	1.24	India	3,302	3,255	CIDI	2008	50+	WHO SAGE - India		3	3[b]	3
Avenevoli et al. (2015)[a]	1.65	U.S.	843	809	WMH-CIDI	2002	13	National Comorbidity Survey-Adolescent Supplement (NCS- A)	1	2	1	3
Avenevoli et al. (2015)[a]	5.50	U.S.	1,088	1,130	WMH-CIDI	2002	14	NCS- A	1	2	1	3
Avenevoli et al. (2015)[a]	2.57	U.S.	883	1,004	WMH-CIDI	2002	15	NCS- A	1	2	1	3
Avenevoli et al. (2015)[a]	2.44	U.S.	966	1,044	WMH-CIDI	2002	16	NCS- A	1	2	1	3
Avenevoli et al. (2015)[a]	1.93	U.S.	1,173	1,183	WMH-CIDI	2002	17–18	NCS- A	1	2	1	3
Beals et al. (2005)	1.79	U.S.	606	817	CIDI	1998	15–54	The American Indian Service Utilization, Psychiatric Epidemiology, Risk & Protective Factors Project (AI-SUPERPFP)-Southwest Tribe	6	2	2	3
Beals et al. (2005)	1.94	U.S.	778	840	CIDI	1998	15–54	AI-SUPERPFP- Northern Plains Tribe	6	2	2	3
Bijl et al. (1998)	1.90	Netherlands	3,304	3,772	CIDI 1.1	1996	18–64	Netherlands Mental Health Survey & Incidence Study (NEMESIS)		2	2	3
Boyd et al. (2015)	2.55	Bulgaria	2,430	2,888	WMH-CIDI	2005	18+	Bulgaria National Survey of Health & Stress[W]		1	1	4
Boyd et al. (2015)	1.67	Romania	1,092	1,265	WMH-CIDI	2006	18+	Romania Mental Health Survey[W]		1	1	4
Boyd et al. (2015)	2.44	Portugal	1,632	2,217	WMH-CIDI	2009	18+	Portugal National Mental Health Survey[W]		1	1	4
Boyd et al. (2015)	1.96	France	1,329	1,565	WMH-CIDI	2002	18+	ESEMeD[W]		1	1	4
Bromet et al. (2005)[a]	3.58	Ukraine	462	608	WMH-CIDI	2002	18–32	Comorbid Mental Disorders during Periods of Social Disruption (CMDPSD)[W]		2	1	3
Bromet et al. (2005)[a]	2.31	Ukraine	443	728	WMH-CIDI	2002	33–47	CMDPSD[W]		2	1	3
Bromet et al. (2005)[a]	2.07	Ukraine	440	876	WMH-CIDI	2002	48–62	CMDPSD[W]		2	1	3
Bromet et al. (2005)[a]	2.07	Ukraine	447	876	WMH-CIDI	2002	63+	CMDPSD[W]		2	1	3
Bromet et al. (2011)	1.57	Belgium	1,190	1,229	WMH-CIDI	2002	18+	The European Study of the Epidemiology of Mental Disorders (ESEMeD)[W]		2	1	3
Bromet et al. (2011)	1.71	Germany	1,660	1,895	WMH-CIDI	2003	18+	ESEMeD[W]		2	1	3
Bromet et al. (2011)	1.60	Israel	2,380	2,479	WMH-CIDI	2003	21+	Israel National Health Survey[W]		2	1	3
Bromet et al. (2011)	2.54	Italy	2,321	2,391	WMH-CIDI	2002	18+	ESEMeD[W]		2	1	3
Bromet et al. (2011)	2.32	Netherlands	1,032	1,340	WMH-CIDI	2003	18+	ESEMeD[W]		2	1	3
Bromet et al. (2011)	1.96	Colombia	1,700	2,726	WMH-CIDI	2003	18–65	Colombian National Study of Mental Health[W]		2	1	3
Bromet et al. (2011)	2.11	Lebanon	1,297	1,560	WMH-CIDI	2003	18+	Lebanese Evaluation of the Burden of Ailments and Needs of the Nation[W]		2	1	3
Bromet et al. (2011)	2.17	South Africa	1,718	2,597	WMH-CIDI	2004	18+	South Africa Stress and Health Study[W]		2	1	3
Center for Behavioral Health Statistics and Quality (CBHSQ). (2014)	2.86	U.S.	11,363	10,938	WMH-CIDI	2004	12–17	National Survey on Drug Use & Health (NSDUH)	1	3	1	3
CBHSQ (2014)	3.26	U.S.	11,378	11,156	WMH-CIDI	2005	12–17	NSDUH	1	3	1	3
CBHSQ (2014)	3.05	U.S.	11,718	11,153	WMH-CIDI	2006	12–17	NSDUH	1	3	1	3
CBHSQ (2014)	2.80	U.S.	11,524	10,909	WMH-CIDI	2007	12–17	NSDUH	1	3	1	3
CBHSQ (2014)	3.18	U.S.	11,517	11,029	WMH-CIDI	2008	12–17	NSDUH	1	3	1	3
CBHSQ (2014)	2.69	U.S.	11,520	11,106	WMH-CIDI	2009	12–17	NSDUH	1	3	1	3
CBHSQ (2014)	2.93	U.S.	11,140	10,820	WMH-CIDI	2010	12–17	NSDUH	1	3	1	3
CBHSQ (2014)	2.92	U.S.	12,028	11,482	WMH-CIDI	2011	12–17	NSDUH	1	3	1	3
CBHSQ (2014)	2.66	U.S.	1,918	1,807	WMH-CIDI	2012	12	NSDUH	1	3	1	3
CBHSQ (2014)	5.22	U.S.	1,840	1,838	WMH-CIDI	2012	13	NSDUH	1	3	1	3

(table continues)

Table 1 (*continued*)

Study	OR	Country	NM	NF	Interview	Year	Age	Sample	E	F	M	T
CBHSQ (2014)	3.28	U.S.	1,883	1,872	WMH-CIDI	2012	14	NSDUH	1	3	1	3
CBHSQ (2014)	2.98	U.S.	1,921	1,817	WMH-CIDI	2012	15	NSDUH	1	3	1	3
CBHSQ (2014)	3.71	U.S.	1,937	1,878	WMH-CIDI	2012	16	NSDUH	1	3	1	3
CBHSQ (2014)	2.66	U.S.	1,877	1,885	WMH-CIDI	2012	17	NSDUH	1	3	1	3
CBHSQ (2014)	2.14	U.S.	1,824	1,713	WMH-CIDI	2013	12	NSDUH	1	3	1	3
CBHSQ (2014)	3.76	U.S.	1,963	1,849	WMH-CIDI	2013	13	NSDUH	1	3	1	3
CBHSQ (2014)	4.95	U.S.	2,026	1,865	WMH-CIDI	2013	14	NSDUH	1	3	1	3
CBHSQ (2014)	3.98	U.S.	1,882	1,868	WMH-CIDI	2013	15	NSDUH	1	3	1	3
CBHSQ (2014)	2.92	U.S.	1,940	1,890	WMH-CIDI	2013	16	NSDUH	1	3	1	3
CBHSQ (2014)	3.12	U.S.	1,914	1,760	WMH-CIDI	2013	17	NSDUH	1	3	1	3
CBHSQ (2014)	1.99	U.S.	10,774	11,755	WMH-CIDI	2012	18–25	NSDUH	1	3	1	3
CBHSQ (2014)	1.89	U.S.	1,469	1,593	WMH-CIDI	2012	26–29	NSDUH	1	3	1	3
CBHSQ (2014)	2.13	U.S.	1,559	1,863	WMH-CIDI	2012	30–34	NSDUH	1	3	1	3
CBHSQ (2014)	1.39	U.S.	1,365	1,537	WMH-CIDI	2012	35–39	NSDUH	1	3	1	3
CBHSQ (2014)	1.69	U.S.	1,394	1,615	WMH-CIDI	2012	40–44	NSDUH	1	3	1	3
CBHSQ (2014)	1.73	U.S.	1,428	1,737	WMH-CIDI	2012	45–49	NSDUH	1	3	1	3
CBHSQ (2014)	1.35	U.S.	816	1,013	WMH-CIDI	2012	50–54	NSDUH	1	3	1	3
CBHSQ (2014)	1.45	U.S.	722	877	WMH-CIDI	2012	55–59	NSDUH	1	3	1	3
CBHSQ (2014)	1.76	U.S.	598	736	WMH-CIDI	2012	60–65	NSDUH	1	3	1	3
CBHSQ (2014)	2.09	U.S.	1,360	1,625	WMH-CIDI	2012	65+	NSDUH	1	3	1	3
CBHSQ (2014)	2.09	U.S.	10,671	11,543	WMH-CIDI	2013	18–25	NSDUH	1	3	1	3
CBHSQ (2014)	1.80	U.S.	1,376	1,603	WMH-CIDI	2013	26–29	NSDUH	1	3	1	3
CBHSQ (2014)	1.29	U.S.	1,529	1,802	WMH-CIDI	2013	30–34	NSDUH	1	3	1	3
CBHSQ (2014)	1.67	U.S.	1,317	1,562	WMH-CIDI	2013	35–39	NSDUH	1	3	1	3
CBHSQ (2014)	1.63	U.S.	1,437	1,671	WMH-CIDI	2013	40–44	NSDUH	1	3	1	3
CBHSQ (2014)	1.34	U.S.	1,440	1,631	WMH-CIDI	2013	45–49	NSDUH	1	3	1	3
CBHSQ (2014)	1.47	U.S.	837	951	WMH-CIDI	2013	50–54	NSDUH	1	3	1	3
CBHSQ (2014)	1.61	U.S.	711	909	WMH-CIDI	2013	55–59	NSDUH	1	3	1	3
CBHSQ (2014)	1.42	U.S.	674	719	WMH-CIDI	2013	60–65	NSDUH	1	3	1	3
CBHSQ (2014)	3.55	U.S.	1,302	1,659	WMH-CIDI	2013	65+	NSDUH	1	3	1	3
CBHSQ (2014)	1.90	U.S.	21,408	24,366	WMH-CIDI	2005	18+	NSDUH	1	3	1	3
CBHSQ (2014)	1.75	U.S.	20,995	23,936	WMH-CIDI	2006	18+	NSDUH	1	3	1	3
CBHSQ (2014)	1.91	U.S.	21,272	24,165	WMH-CIDI	2007	18+	NSDUH	1	3	1	3
CBHSQ (2014)	1.88	U.S.	21,602	24,588	WMH-CIDI	2008	18+	NSDUH	1	3	1	3
CBHSQ (2014)	1.77	U.S.	21,755	24,319	WMH-CIDI	2009	18+	NSDUH	1	3	1	3
CBHSQ (2014)	1.73	U.S.	21,697	24,147	WMH-CIDI	2010	18+	NSDUH	1	3	1	3
CBHSQ (2014)	1.84	U.S.	21,750	24,489	WMH-CIDI	2011	18+	NSDUH	1	3	1	3
Chan Chee et al. (2011)	2.28	France	529	596	CIDI-SF	2005	15–19	French Barometer Study		2	1	3
Chan Chee et al. (2011)	1.98	France	1,813	2,417	CIDI-SF	2005	20–34	French Barometer Study		2	1	3
Chan Chee et al. (2011)	2.46	France	2,655	3,690	CIDI-SF	2005	35–54	French Barometer Study		2	1	3
Chan Chee et al. (2011)	1.95	France	2,081	3,102	CIDI-SF	2005	55–75	French Barometer Study		2	1	3
Chan Chee et al. (2011)	2.02	France	878	1,060	CIDI-SF	2010	20–34	French Barometer Study		2	1	3
Chan Chee et al. (2011)	1.52	France	1,433	1,725	CIDI-SF	2010	35–54	French Barometer Study		2	1	3
Chan Chee et al. (2011)	2.65	France	1,122	1,527	CIDI-SF	2010	55–75	French Barometer Study		2	1	3
Cho et al. (2010)[a]	2.40	South Korea	560	683	K-CIDI 2.1	2007	18–29	Korean National Epidemiologic Catchment Area Study- Replication (KECA-R)		2	1	3
Cho et al. (2010)[a]	2.55	South Korea	688	1,136	K-CIDI 2.1	2007	30–39	KECA-R		2	1	3
Cho et al. (2010)[a]	1.26	South Korea	690	1,065	K-CIDI 2.1	2007	40–49	KECA-R		2	1	3
Cho et al. (2010)[a]	1.85	South Korea	431	683	K-CIDI 2.1	2007	50–59	KECA-R		2	1	3
Cho et al. (2007)	3.18	South Korea	3,524	2,751	K-CIDI 2.1	2001	18–64	KECA		2	1	3
Coyne and Marcus (2006)	2.15	U.S.	11,612	14,903	CIDI-SF	1999	18+	National Health Interview Survey	2	2	1	3
Coyne and Marcus (2006)	2.29	U.S.	1,590	2,696	CIDI-SF	1999	18+	National Health Interview Survey	3	2	1	3
Danielson et al. (2005)	1.98	U.S.	2,020	2,003	NSA interview	1995	12–17	National Survey of Adolescents (NSA)	1	2	1	3
de Graaf et al. (2012)	1.57	Netherlands	2,977	3,669	WMH-CIDI	2008	18–64	NEMESIS II		2	1	3
Farbstein et al. (2010)	2.55	Israel	497	460	DAWBA	2005	14–17	Israel Survey of Mental Health among Adolescents (ISMEHA)		3	1	2
Ford et al. (2003)[a]	1.51	Great Britain	1,284	1,340	DAWBA	1999	13–15	British Child & Adolescent Mental Health Survey		2	1	0
Gabilondo et al. (2010)	2.83	Spain	733	834	WMH-CIDI	2002	18–34	ESEMeD[W]		2	1	3
Gabilondo et al. (2010)	1.82	Spain	622	809	WMH-CIDI	2002	35–49	ESEMeD[W]		2	1	3
Gabilondo et al. (2010)	2.83	Spain	437	587	WMH-CIDI	2002	50–64	ESEMeD[W]		2	1	3
Gabilondo et al. (2010)	6.29	Spain	629	822	WMH-CIDI	2002	65+	ESEMeD[W]		2	1	3
Gavin et al. (2009)	1.92	U.S.	1,217	2,217	WMH-CIDI	2002	18+	National Survey of American Life	3	2	1	4
Goodwin and Gotlib (2004)	1.75	U.S.	1,492	1,540	CIDI-SF	1996	25–74	MIDUS (Midlife in the US)	1	1	2	3
Graham et al. (2007)	2.06	Canada	6,214	7,878	CIDI	2005	18–76	Gender Alcohol and Culture: An International Study (GENACIS)		2	1	3

Table 1 (*continued*)

Study	OR	Country	NM	NF	Interview	Year	Age	Sample	E	F	M	T
Grant (1995)	1.31	U.S.	21,431	21,431	AUDADIS	1992	18+	National Longitudinal Alcohol Epidemiologic Survey (NLAES)	1	2	1	4
Gureje et al. (2010)[a]	.83	Nigeria	1,595	1,580	WMH-CIDI	2003	18–34	Nigerian Survey of Mental Health & Wellbeing (NSMHW)[W]		2	1	3
Gureje et al. (2010)[a]	1.84	Nigeria	809	822	WMH-CIDI	2003	35–49	NSMHW[W]		2	1	3
Gureje et al. (2010)[a]	1.28	Nigeria	503	601	WMH-CIDI	2003	50–64	NSMHW[W]		2	1	3
Haarasilta et al. (2001)	1.36	Finland	437	505	CIDI-SF	1996	15–24	Finnish Health Care Survey (FINHCS)		2	2	3
Hasin et al. (2005)[a]	1.93	U.S.	2,410	2,789	AUDADIS	2002	18–24	National Epidemiologic Survey on Alcohol & Related Conditions (NESARC)	1	2	1	3
Hasin et al. (2005)[a]	2.08	U.S.	1,407	2,060	AUDADIS	2002	25–29	NESARC	1	2	1	3
Hasin et al. (2005)[a]	3.28	U.S.	1,834	2,458	AUDADIS	2002	30–34	NESARC	1	2	1	3
Hasin et al. (2005)[a]	1.65	U.S.	1,989	2,661	AUDADIS	2002	35–39	NESARC	1	2	1	3
Hasin et al. (2005)[a]	2.36	U.S.	2,034	2,406	AUDADIS	2002	40–44	NESARC	1	2	1	3
Hasin et al. (2005)[a]	1.84	U.S.	1,876	2,142	AUDADIS	2002	45–49	NESARC	1	2	1	3
Hasin et al. (2005)[a]	2.53	U.S.	1,603	2,004	AUDADIS	2002	50–54	NESARC	1	2	1	3
Hasin et al. (2005)[a]	1.76	U.S.	1,236	1,611	AUDADIS	2002	55–59	NESARC	1	2	1	3
Hasin et al. (2005)[a]	1.41	U.S.	1,025	1,343	AUDADIS	2002	60–64	NESARC	1	2	1	3
Hasin et al. (2005)[a]	2.38	U.S.	927	1,255	AUDADIS	2002	65–69	NESARC	1	2	1	3
Hasin et al. (2005)[a]	1.74	U.S.	2,177	3,846	AUDADIS	2002	70+	NESARC	1	2	1	3
Instituto Nacional de Salud Publica (2014)	6.02	Mexico	1,028	1,176	CIDI	2010	50+	WHO SAGE- Mexico		3	3[b]	3
Joe et al. (2009)[a]	.98	U.S.	563	607	WMH-CIDI	2002	13–17	National Survey of American Life	3	3	1	3
Kessler et al. (2010)	2.03	U.S.	1,375	1,658	WMH-CIDI	2002	18–34	National Comorbidity Survey- Replication (NCS- R)[W]	1	2	1	3
Kessler et al. (2010)	1.59	U.S.	1,343	1,522	WMH-CIDI	2002	35–49	NCS-R[W]	1	2	1	3
Kessler et al. (2010)	1.54	U.S.	854	1,068	WMH-CIDI	2002	50–64	NCS-R[W]	1	2	1	3
Kessler et al. (2010)	3.16	U.S.	567	894	WMH-CIDI	2002	65+	NCS-R[W]	1	2	1	3
Kessler et al. (1993)	1.86	U.S.	838	927	CIDI	1991	15–24	National Comorbidity Survey (NCS)	1	1	2	3
Kessler et al. (1993)	1.53	U.S.	1,246	1,378	CIDI	1991	25–34	NCS	1	1	2	3
Kessler et al. (1993)	1.65	U.S.	1,065	1,178	CIDI	1991	35–44	NCS	1	1	2	3
Kessler et al. (1993)	2.97	U.S.	696	770	CIDI	1991	45–54	NCS	1	1	2	3
Kiejna et al. (2015)	1.82	Poland	1,465	1,430	WMH-CIDI	2011	18–29	Epidemiology of Mental Health and Access to Care Survey (EZOP)[W]		1	1	4
Kiejna et al. (2015)	1.76	Poland	1,069	1,050	WMH-CIDI	2011	30–39	EZOP[W]		1	1	4
Kiejna et al. (2015)	2.51	Poland	938	936	WMH-CIDI	2011	40–49	EZOP[W]		1	1	4
Kiejna et al. (2015)	2.37	Poland	1,411	1,782	WMH-CIDI	2011	50–64	EZOP[W]		1	1	4
Kim et al. (2015)[a]	2.36	South Korea	691	636	K-CIDI	2011	18–29	KECA-2011	1	1	1	3
Kim et al. (2015)[a]	2.02	South Korea	655	645	K-CIDI	2011	30–39	KECA-2011	1	1	1	3
Kim et al. (2015)[a]	2.50	South Korea	687	682	K-CIDI	2011	40–49	KECA-2011	1	1	1	3
Kim et al. (2015)[a]	2.01	South Korea	542	553	K-CIDI	2011	50–59	KECA-2011	1	1	1	3
Kim et al. (2015)[a]	3.40	South Korea	315	351	K-CIDI	2011	60–69	KECA-2011	1	1	1	3
Kim et al. (2015)[a]	3.65	South Korea	112	149	K-CIDI	2011	70+	KECA-2011	1	1	1	3
Lara et al. (2015)	1.80	Spain	435	523	CIDI	2012	18–49	Collaborative Research on Aging in Europe (COURAGE)- Spain		2	1	3
Lara et al. (2015)	1.95	Spain	829	931	CIDI	2012	50–64	COURAGE- Spain		2	1	3
Lara et al. (2015)	4.17	Spain	814	1051	CIDI	2012	65+	COURAGE- Spain		2	1	3
Lepine et al. (1997)	1.75	Belgium	4,032	4,044	MINI	1995	18+	DEPRES (Depression Research in European Society)		2	1	2
Lepine et al. (1997)	2.22	France	7,162	7,355	MINI	1995	15+	DEPRES		2	1	2
Lepine et al. (1997)	1.58	Germany	7,798	8,386	MINI	1995	14+	DEPRES		2	1	2
Lepine et al. (1997)	1.97	Netherlands	4,224	3,587	MINI	1995	16+	DEPRES		2	1	2
Lepine et al. (1997)	2.12	Spain	8,063	8,069	MINI	1995	15+	DEPRES		2	1	2
Lepine et al. (1997)	1.41	U.K.	7,155	8,588	MINI	1995	16+	DEPRES		2	1	2
Lorenzo-Blanco et al. (2013)	1.77	U.S.	1,127	1,427	WMH-CIDI	2003	18+	National Latino & Asian American Study (NLAAS)	5	1	1	3
Maske et al. (2016)[a]	3.69	Germany	371	413	M-CIDI	2011	18–34	Germany Health Interview and Examination Survey for Adults- mental health module (DEGS1-MH)		2	1	3
Maske et al. (2016)[a]	4.33	Germany	291	332	M-CIDI	2011	35–44	DEGS1-MH		2	1	3
Maske et al. (2016)[a]	1.41	Germany	414	524	M-CIDI	2011	45–54	DEGS1-MH		2	1	3
Maske et al. (2016)[a]	1.21	Germany	399	444	M-CIDI	2011	55–64	DEGS1-MH		2	1	3
Maske et al. (2016)[a]	3.10	Germany	628	592	M-CIDI	2011	65–79	DEGS1-MH		2	1	3
McMartin et al. (2013)[a]	2.82	Canada	3,228	3,078	CIDI-SF	2001	12–14	Canadian Community Health Survey (CCHS)		2	1	3

(*table continues*)

Table 1 (*continued*)

Study	OR	Country	NM	NF	Interview	Year	Age	Sample	E	F	M	T
McMartin et al. (2013)[a]	3.22	Canada	5,425	5,380	CIDI-SF	2001	15–19	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	1.76	Canada	3,350	3,997	CIDI-SF	2001	20–24	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	1.88	Canada	3,914	4,586	CIDI-SF	2001	25–29	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	1.87	Canada	4,644	5,306	CIDI-SF	2001	30–34	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	1.85	Canada	5,687	6,429	CIDI-SF	2001	35–39	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	2.17	Canada	6,073	6,458	CIDI-SF	2001	40–44	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	2.06	Canada	5,368	5,647	CIDI-SF	2001	45–49	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	1.45	Canada	4,759	5,184	CIDI-SF	2001	50–54	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	2.23	Canada	3,867	4,226	CIDI-SF	2001	55–59	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	1.62	Canada	3,136	3,782	CIDI-SF	2001	60–64	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	1.72	Canada	2,988	3,625	CIDI-SF	2001	65–69	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	1.91	Canada	2,607	3,572	CIDI-SF	2001	70–74	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	2.27	Canada	1,914	3,140	CIDI-SF	2001	75–79	CCHS 1.1		2	1	3
McMartin et al. (2013)[a]	1.54	Canada	1,783	3,746	CIDI-SF	2001	80+	CCHS 1.1		2	1	3
Merikangas et al. (2012)	2.98	U.S.	2,147	2,003	NIMH DISC-IV	2003	12–19	National Health & Nutrition Examination Survey (NHANES), 2001–2004	1	2	1	3
Mohammadi et al. (2005)	2.84	Iran	12,660	12,530	SADS	2001	18+	The national plan for epidemiologic study of psychiatric disorders in Iran		2	1	4
Mommersteeg et al. (2013)[a]	1.24	Burkina Faso	2,272	2,551	WHO WHS	2002	18+	World Health Organization (WHO) World Health Survey (WHS)		2	1	3
Mommersteeg et al. (2013)[a]	1.42	Chad	2,196	2,447	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.70	Comoros	787	972	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.64	Congo	1,167	1,326	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.11	Ethiopia	2,391	2,544	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.67	Ghana	1,773	2,153	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.66	Ivory Coast	1,818	1,361	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.03	Kenya	1,868	2,541	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.17	Malawi	2,197	3,033	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.00	Mali	2,354	1,711	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.13	Mauritania	1,465	2,308	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.41	Mauritius	1,872	2,016	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.25	Namibia	1,721	2,524	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.01	Senegal	1,641	1,515	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	.93	Swaziland	1,417	1,667	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.29	Tunisia	2,344	2,725	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.77	Zimbabwe	1,489	2,600	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.57	Brazil	2,188	2,812	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.94	Dominican Republic	2,104	2,430	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.85	Ecuador	2,051	2,602	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.48	Mexico	16,377	22,368	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	3.39	Paraguay	2,353	2,789	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	3.25	Uruguay	1,449	1,530	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.82	Bangladesh	2,584	2,966	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.01	China	1,954	2,039	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.21	Georgia	1,165	1,590	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.30	India	4,849	5,144	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.65	Kazakhstan	1,544	2,951	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.63	Lao PDR	2,295	2,594	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.37	Malaysia	2,673	3,366	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.01	Myanmar	2,551	3,335	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.52	Nepal	3,698	4,990	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.34	Pakistan	3,565	2,810	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.49	Philippines	4,659	5,416	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	.83	Sri Lanka	3,136	3,596	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	.84	UAE	617	563	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	7.04	Vietnam	1,572	1,919	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	3.82	Bosnia & Herzegovina	434	594	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.42	Croatia	401	589	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.46	Czech Republic	419	516	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.75	Estonia	366	645	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.57	Hungary	591	828	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.92	Russia	1,592	2,828	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	3.08	Slovakia	952	1,530	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	1.70	Spain	2,623	3,740	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
Mommersteeg et al. (2013)[a]	2.03	Ukraine	1,001	1,844	WHO WHS	2002	18+	WHO World Health Survey		2	1	3
National Research Institute of Public Health, Russian Academy of Medical Science (2013)	3.02	Russia	1,520	2,376	CIDI	2009	50+	WHO SAGE- Russia		3	3[b]	3
Patten (2005)[a]	2.74	Canada	1,432	1,434	WMH-CIDI	2002	15–19	Canadian Community Health Survey (CCHS) 1.2		2	1	3

Table 1 (*continued*)

Study	OR	Country	NM	NF	Interview	Year	Age	Sample	E	F	M	T
Patten (2005)[a]	1.48	Canada	1,305	1,502	WMH-CIDI	2002	20–24	CCHS 1.2		2	1	3
Patten (2005)[a]	2.20	Canada	1,074	1,468	WMH-CIDI	2002	25–29	CCHS 1.2		2	1	3
Patten (2005)[a]	1.64	Canada	1,498	1,731	WMH-CIDI	2002	30–34	CCHS 1.2		2	1	3
Patten (2005)[a]	1.51	Canada	1,669	1,860	WMH-CIDI	2002	35–39	CCHS 1.2		2	1	3
Patten (2005)[a]	1.80	Canada	1,775	1,738	WMH-CIDI	2002	40–44	CCHS 1.2		2	1	3
Patten (2005)[a]	1.67	Canada	1,443	1,464	WMH-CIDI	2002	45–49	CCHS 1.2		2	1	3
Patten (2005)[a]	2.69	Canada	1,282	1,517	WMH-CIDI	2002	50–54	CCHS 1.2		2	1	3
Patten (2005)[a]	1.34	Canada	1,264	1,465	WMH-CIDI	2002	55–59	CCHS 1.2		2	1	3
Patten (2005)[a]	.89	Canada	1,092	1,235	WMH-CIDI	2002	60–64	CCHS 1.2		2	1	3
Patten (2005)[a]	1.37	Canada	889	1,184	WMH-CIDI	2002	65–69	CCHS 1.2		2	1	3
Patten (2005)[a]	1.13	Canada	834	1,206	WMH-CIDI	2002	70–74	CCHS 1.2		2	1	3
Patten (2005)[a]	.97	Canada	599	1,042	WMH-CIDI	2002	75–79	CCHS 1.2		2	1	3
Patten (2005)[a]	.22	Canada	617	1,365	WMH-CIDI	2002	80+	CCHS 1.2		2	1	3
Patten (2005)[a]	3.03	Canada	967	1,057	WMH-CIDI	2012	15–19	CCHS- MH (Mental Health)		2	1	3
Patten (2005)[a]	1.17	Canada	923	1,066	WMH-CIDI	2012	20–24	CCHS- MH		2	1	3
Patten (2005)[a]	2.09	Canada	737	880	WMH-CIDI	2012	25–29	CCHS- MH		2	1	3
Patten (2005)[a]	2.00	Canada	808	1,061	WMH-CIDI	2012	30–34	CCHS- MH		2	1	3
Patten (2005)[a]	1.52	Canada	827	902	WMH-CIDI	2012	35–39	CCHS- MH		2	1	3
Patten (2005)[a]	1.41	Canada	877	814	WMH-CIDI	2012	40–44	CCHS- MH		2	1	3
Patten (2005)[a]	1.37	Canada	795	875	WMH-CIDI	2012	45–49	CCHS- MH		2	1	3
Patten (2005)[a]	2.04	Canada	904	1,052	WMH-CIDI	2012	50–54	CCHS- MH		2	1	3
Patten (2005)[a]	1.86	Canada	1,016	1,229	WMH-CIDI	2012	55–59	CCHS- MH		2	1	3
Patten (2005)[a]	1.57	Canada	975	1,231	WMH-CIDI	2012	60–64	CCHS- MH		2	1	3
Patten (2005)[a]	2.31	Canada	847	1,071	WMH-CIDI	2012	65–69	CCHS- MH		2	1	3
Patten (2005)[a]	.39	Canada	639	793	WMH-CIDI	2012	70–74	CCHS- MH		2	1	3
Patten (2005)[a]	1.13	Canada	466	718	WMH-CIDI	2012	75–79	CCHS- MH		2	1	3
Patten (2005)[a]	5.07	Canada	559	1,024	WMH-CIDI	2012	80+	CCHS- MH		2	1	3
Peltzer & Phaswana-Mafuya (2013)	1.08	South Africa	1,638	2,202	WMH-CIDI	2008	50+	WHO SAGE- South Africa		2	3[b]	3
Pirkola et al. (2005)[a]	2.26	Finland	983	1,094	CIDI	2001	30–44	Health, 2000		2	1	3
Pirkola et al. (2005)	1.62	Finland	768	812	CIDI	2001	45–54	Health, 2000		2	1	3
Pirkola et al. (2005)	2.10	Finland	486	563	CIDI	2001	55–64	Health, 2000		2	1	3
Pirkola et al. (2005)	2.49	Finland	511	788	CIDI	2001	65+	Health, 2000		2	1	3
Rafful et al. (2012)	1.65	Mexico	842	1,218	WMH-CIDI	2002	18–29	Mexico National Comorbidity Survey (M-NCS)[W]		1	1	4
Rafful et al. (2012)	2.40	Mexico	852	1,384	WMH-CIDI	2002	30–44	M-NCS[W]		1	1	4
Rafful et al. (2012)	2.20	Mexico	591	895	WMH-CIDI	2002	45–65	M-NCS[W]		1	1	4
Rapsey et al. (2015)[a]	1.91	Iraq	2,091	2,241	WMH-CIDI	2007	18–96	Iraq Mental Health Survey [W]		3	1	4
Sandanger et al. (2007)	1.55	Norway	803	888	CIDI 1.2	2001	18+	—		1	1	3
Scott et al. (2010)[a]	1.61	New Zealand	713	822	WMH-CIDI	2004	16–24	New Zealand Mental Health Survey (NZMHS)[W]		2	1	3
Scott et al. (2010)[a]	2.10	New Zealand	1,000	1,414	WMH-CIDI	2004	25–34	NZMHS[W]		2	1	3
Scott et al. (2010)[a]	1.23	New Zealand	1,231	1,659	WMH-CIDI	2004	35–44	NZMHS[W]		2	1	3
Scott et al. (2010)[a]	1.70	New Zealand	1,023	1,222	WMH-CIDI	2004	45–54	NZMHS[W]		2	1	3
Scott et al. (2010)[a]	3.28	New Zealand	730	934	WMH-CIDI	2004	55–64	NZMHS[W]		2	1	3
Scott et al. (2010)[a]	3.31	New Zealand	937	1,307	WMH-CIDI	2004	65+	NZMHS[W]		2	1	3
Shah et al. (2011)	1.84	U.S.	1,107	1,423	DIS	1991	17–39	NHANES- III	3	2	2	4
Shah et al. (2011)	2.34	U.S.	948	1,236	DIS	1991	17–39	NHANES- III	2	2	2	4
Shah et al. (2011)	2.16	U.S.	1,281	1,329	DIS	1991	17–39	NHANES- III	5	2	2	4
Shanghai Municipal Center for Disease Control and Prevention (2012)	1.58	China	6,409	6,466	CIDI	2009	50+	WHO SAGE- China		3	3[b]	3
Skapinakis et al. (2013)	1.78	Greece	2,427	2,467	CIS-R	2010	18–70	Psychiatric Morbidity Survey		2	3[b]	1
Slade et al. (2009)[a]	2.92	Australia	681	790	WMH-CIDI	2007	16–24	National Survey of Mental Health and Well-Being (NSMHWB)[W]		2	3[b]	3
Slade et al. (2009)[a]	1.57	Australia	516	774	WMH-CIDI	2007	25–34	NSMHWB[W]		2	3[b]	3
Slade et al. (2009)[a]	1.17	Australia	756	882	WMH-CIDI	2007	35–44	NSMHWB[W]		2	3[b]	3
Slade et al. (2009)[a]	1.34	Australia	566	698	WMH-CIDI	2007	45–54	NSMHWB[W]		2	3[b]	3
Slade et al. (2009)[a]	2.80	Australia	604	669	WMH-CIDI	2007	55–64	NSMHWB[W]		2	3[b]	3
Slade et al. (2009)[a]	1.75	Australia	904	1,001	WMH-CIDI	2007	65–85	NSMHWB[W]		2	3[b]	3
Spiers et al. (2012)	1.60	England	4,300	4,318	CIS-R	1993	16–64	National Psychiatric Morbidity Surveys (NPMS)		2	3[b]	1
Spiers et al. (2012)	1.17	England	3,606	3,622	CIS-R	2000	16–71	NPMS		2	3[b]	1
Spiers et al. (2012)	1.50	England	3,454	3,553	CIS-R	2007	16–78	NPMS		2	3[b]	1
Subramaniam et al. (2013)[a]	1.43	Singapore	1,149	1,144	WMH-CIDI	2010	18–34	Singapore Mental Health Study		2	1	3
Subramaniam et al. (2013)[a]	1.00	Singapore	1,162	1,197	WMH-CIDI	2010	35–49	Singapore Mental Health Study		2	1	3
Subramaniam et al. (2013)[a]	3.02	Singapore	978	976	WMH-CIDI	2010	50–89	Singapore Mental Health Study		2	1	3
Suttajit et al. (2012)[a]	2.26	Thailand	1,886	1,786	MINI	2008	15–24	Thai National Mental Health survey		2	1	4

Table 1 (*continued*)

Study	OR	Country	NM	NF	Interview	Year	Age	Sample	E	F	M	T
Suttajit et al. (2012)[a]	1.52	Thailand	1,831	1,908	MINI	2008	25–34	Thai National Mental Health survey		2	1	4
Suttajit et al. (2012)[a]	1.25	Thailand	1,904	1,982	MINI	2008	35–44	Thai National Mental Health survey		2	1	4
Suttajit et al. (2012)[a]	1.57	Thailand	2,875	2,968	MINI	2008	45–59	Thai National Mental Health survey		2	1	4
Takeuchi et al. (2007)[a]	.81	U.S.	998	1,097	WMH-CIDI	2003	18+	National Latino & Asian American Study (NLAAS)	4	2	1	3
Toussaint et al. (2008)[a]	2.57	U.S.	563	709	CIDI-SF	1998	18+	—	1	1	2	3
University of Ghana Medical School (2013)	1.72	Ghana	2,241	2,041	CIDI	2008	50+	WHO SAGE- Ghana		3	3[b]	3
Vicente et al. (2006)	2.11	Chile	1,281	1,697	CIDI	1996	15+	Chilean Psychiatric Prevalence Study		2	2	3
Volken (2013)[a]	1.25	Switzerland	7,919	9,936	CIDI- SF	2007	15+	Swiss Health Survey		2	1	3
Wade et al. (2002)	1.93	Canada	643	677	CIDI-SF	1995	12–19	National Population Health Survey (NPHS)		1	2	3
Wittchen et al. (2000)	2.17	Germany	1,913	2,268	CIDI	1998	18–65	German National Health Interview and Examination Survey		2	1	3
Zhao et al. (2006)	2.21	Canada	12,991	14,399	CIDI-SF	1997	20–39	NPHS		1	2	3
Zhao et al. (2006)	1.91	Canada	12,386	13,296	CIDI-SF	1997	40–64	NPHS		1	2	3
Zinzow et al. (2009)[a]	3.24	U.S.	923	885	NSA interview	2005	12–14	National Survey of Adolescents-Replication (NSA-R)	1	2	1	2
Zinzow et al. (2009)[a]	2.79	U.S.	928	868	NSA interview	2005	15–17	NSA-R	1	2	1	2
Zubrick et al. (2016)[a]	4.96	Australia	183	160	DISC-IV	2014	14	Second Australian Child and Adolescent Survey of Health and Well-being		2	1	3
Zubrick et al. (2016)[a]	3.16	Australia	168	141	DISC-IV	2014	15	Second Australian Child and Adolescent Survey of Health and Well-being		2	1	3
Zubrick et al. (2016)[a]	3.55	Australia	353	365	DISC-IV	2014	16	Second Australian Child and Adolescent Survey of Health and Well-being		2	1	3
Zubrick et al. (2016)[a]	2.08	Australia	325	309	DISC-IV	2014	17	Second Australian Child and Adolescent Survey of Health and Well-being		2	1	3

Note. OR = untransformed Odds ratio; NM = *n* men; NF = *n* women; Interview = diagnostic interview used; Year = year of data collection; Age = age or age range (in years); Sample = sample name (not all samples had a name); E = ethnicity (only applies to U.S. samples): 1 = mixed, 2 ≥ 85% White, 3 ≥ 85% African American, 4 ≥ 85% Asian American, 5 ≥ 85% Hispanic, 6 ≥ 85% Native American; F = focus of article: 1 = gender, 2 = depression, 3 = other; M = diagnostic manual or classification system: 1 = *DSM-IV* or *DSM-IV-TR*; 2 = *DSM-III* or *DSM-III-R*; 3 = *ICD-10*; T = time span during which depression was diagnosed: 0 = current; 1 = 1 month, 2 = 6 months, 3 = 12 months, 4 = lifetime; [a] = received additional data from author; [W] = World Mental Health Survey Sample; [b] = *ICD-10* depressive episode (all other diagnoses were major depressive episodes or major depressive disorders); DIS = Diagnostic Interview Schedule (Robins et al., 1981); CIDI = Composite International Diagnostic Interview (World Health Organization, 1990, 1997; Smeets & Dingemans, 1993); WMH-CIDI = World Mental Health CIDI or CIDI 3.0 (Kessler & Ustun, 2004); CIDI-SF = CIDI- Short Form (Kessler et al., 1998); K-CIDI = Korean version of the CIDI (Cho et al., 2002); DAWBA = Development and Well-Being Assessment (Goodman et al., 2000); AUDADIS = Alcohol Use Disorder and Associated Disabilities Interview Schedule (Grant et al., 2001); MINI = Mini-International Neuropsychiatric Interview (Sheehan et al., 1998); M-CIDI = German version of the CIDI (Wittchen & Pfister, 1997); NIMH DISC-IV = National Institute of Mental Health Diagnostic Interview Schedule for Children Version IV (Shaffer et al., 2000); SADS = Schedule of Affective Disorders and Schizophrenia (Endicott & Spitzer, 1978); WHO WHS = World Health Organization World Health Survey (World Health Organization, 2002); CIS-R = Clinical Interview Schedule- Revised (Lewis et al., 1992).

reported (e.g., 12-month and lifetime), 12-month was the preferred time span given our interest in developmental trends in gender differences in depression. For the major depression meta-analysis, all articles reported *DSM* major depression diagnoses or episodes (96%) or the roughly equivalent *ICD-10* depressive episode (4%). Diagnoses of dysthymia were not included in this meta-analysis.

Nation-Level Economic Indicators

To test the contribution of national wealth and development to gender differences in depression, we used two indicators: a categorical measure of national wealth (low-, middle-, and high-income), and

income inequality. The World Bank classification of income categories was based on gross national income (GNI) per capita from 2003: high income ($9,386 or more), middle income ($766–9,385), and low income ($765 or less). Income categories were obtained from the Human Development Report, 2005 (United Nations Development Programme, 2005). The GINI index was used as a measure of income inequality in a nation. It indicates the extent to which the income distribution deviates from a perfectly equal distribution across income categories, with gender not taken into account (World Bank, 2004). High scores indicate greater inequality. See supplemental material Table 1 for a list of each country and its associated indicators. For each of the economic and gender equity indicators,

Table 2

Studies of Gender Differences in Depression Symptoms

Study	d	Country	NM	NF	Measure	Year	Age	Sample	E	F
Aalto et al. (2012)[a]	.27	Finland	326	371	BDI	2001	30–34	Health 2000		2
Aalto et al. (2012)[a]	.29	Finland	353	388	BDI	2001	35–39	Health 2000		2
Aalto et al. (2012)[a]	.14	Finland	341	381	BDI	2001	40–44	Health 2000		2
Aalto et al. (2012)[a]	.10	Finland	379	435	BDI	2001	45–49	Health 2000		2
Aalto et al. (2012)[a]	.16	Finland	412	412	BDI	2001	50–54	Health 2000		2
Aalto et al. (2012)[a]	.37	Finland	263	304	BDI	2001	55–59	Health 2000		2
Aalto et al. (2012)[a]	.19	Finland	244	289	BDI	2001	60–64	Health 2000		2
Aalto et al. (2012)[a]	.31	Finland	178	231	BDI	2001	65–69	Health 2000		2
Aalto et al. (2012)[a]	.22	Finland	348	656	BDI	2001	70+	Health 2000		2
Abebe et al. (2016)[a]	.47	Norway	12,867	13,146	DMI	2012	13	Ungdata		2
Abebe et al. (2016)[a]	.64	Norway	11,018	11,184	DMI	2012	14	Ungdata		2
Abebe et al. (2016)[a]	.70	Norway	12,624	12,369	DMI	2012	15	Ungdata		2
Abebe et al. (2016)[a]	.69	Norway	8,266	7,811	DMI	2012	16	Ungdata		2
Almqvist et al. (1999)	−.11	Finland	2,880	2,805	CDI	1990	8–9	Finnish Nationwide, 1981 Birth Cohort Study		2
Andersen et al. (2009)[a]	.14	Denmark	1,701	2,066	MDI	2000	40–49	Danish Longitudinal Study on Work, Unemployment & Health		2
Andersen et al. (2009)[a]	.14	Denmark	1,695	1,699	MDI	2000	50–56	Danish Longitudinal Study on Work, Unemployment & Health		2
Belanger et al. (2011)[a]	.37	Switzerland	481	355	DTS	2002	16	Swiss Multicenter Adolescent Survey on Health (SMASH)		2
Belanger et al. (2011)[a]	.50	Switzerland	996	976	DTS	2002	17	SMASH		2
Belanger et al. (2011)[a]	.31	Switzerland	1,135	1,072	DTS	2002	18	SMASH		2
Belanger et al. (2011)[a]	.41	Switzerland	782	592	DTS	2002	19	SMASH		2
Belanger et al. (2011)[a]	.08	Switzerland	503	304	DTS	2002	20	SMASH		2
Bracke (1998)	.32	Belgium	2,907	3,204	HDL-D	1992	16+	Panel Study of Belgian Households		1
Bushman et al. (2012)[a]	.08	U.S.	251	549	CES-D	2011	18–90	—	1	2
Cardozo et al. (2005)[a]	.23	Afghanistan	240	357	SCL-D	2002	15+	Mental Health in Afghanistan Survey		1
Cater et al. (2015)[a]	.36	Sweden	203	200	HADS-D	2011	20	Resume Project		1
Cater et al. (2015)[a]	.32	Sweden	215	265	HADS-D	2011	21	Resume Project		1
Cater et al. (2015)[a]	.31	Sweden	253	252	HADS-D	2011	22	Resume Project		1
Cater et al. (2015)[a]	.32	Sweden	265	300	HADS-D	2011	23	Resume Project		1
Cater et al. (2015)[a]	.44	Sweden	250	297	HADS-D	2011	24	Resume Project		1
Chan et al. (2011)	.16	Singapore	759	786	CES-D	2009	60–64	Social Isolation, Health, and Lifestyles Survey (SIHLS)		1
Chan et al. (2011)	.22	Singapore	908	1,039	CES-D	2009	65–74	SIHLS		1
Chan et al. (2011)	.28	Singapore	411	586	CES-D	2009	75+	SIHLS		1
Clark et al. (2013)[a]	.17	New Zealand	762	914	RADS-SF	2013	13	Youth'12		3
Clark et al. (2013)[a]	.33	New Zealand	868	965	RADS-SF	2013	14	Youth'12		3
Clark et al. (2013)[a]	.33	New Zealand	742	941	RADS-SF	2013	15	Youth'12		3
Clark et al. (2013)[a]	.29	New Zealand	686	831	RADS-SF	2013	16	Youth'12		3
Clark et al. (2013)[a]	.32	New Zealand	467	652	RADS-SF	2013	17	Youth'12		3
Collins et al. (2009)[a]	.30	Taiwan	2,534	2,176	CES-D	1996	50+	Survey of Health & Living Status of the Near Elderly & Elderly		2
Crimmins et al. (2011)[a]	.73	Austria	777	1,072	EURO-D	2005	50+	Survey of Health, Aging, & Retirement (SHARE)		1
Crimmins et al. (2011)[a]	1.04	Belgium	1,715	1,934	EURO-D	2005	50+	SHARE		1
Crimmins et al. (2011)[a]	.54	Denmark	757	858	EURO-D	2005	50+	SHARE		1
Crimmins et al. (2011)[a]	.36	France	1,367	1,671	EURO-D	2005	50+	SHARE		1
Crimmins et al. (2011)[a]	.30	Germany	1,370	1,571	EURO-D	2005	50+	SHARE		1
Crimmins et al. (2011)[a]	1.04	Greece	1,241	1,428	EURO-D	2005	50+	SHARE		1
Crimmins et al. (2011)[a]	.36	Italy	1,126	1,382	EURO-D	2005	50+	SHARE		1
Crimmins et al. (2011)[a]	.56	Netherlands	1,348	1,517	EURO-D	2005	50+	SHARE		1
Crimmins et al. (2011)[a]	.57	Spain	989	1,364	EURO-D	2005	50+	SHARE		1
Crimmins et al. (2011)[a]	.96	Sweden	1,407	1,590	EURO-D	2005	50+	SHARE		1
Crimmins et al. (2011)[a]	.66	Switzerland	455	505	EURO-D	2005	50+	SHARE		1
de Wit et al. (2009)[a]	.24	Netherlands	2,632	2,970	MHI-D	2004	18–29	Continuous Survey of Living Conditions (CCSLC)		2
de Wit et al. (2009)[a]	.22	Netherlands	3,206	3,780	MHI-D	2004	30–39	CCSLC		2
de Wit et al. (2009)[a]	.17	Netherlands	3,437	3,727	MHI-D	2004	40–49	CCSLC		2
de Wit et al. (2009)[a]	.17	Netherlands	3,297	3,331	MHI-D	2004	50–59	CCSLC		2
de Wit et al. (2009)[a]	.29	Netherlands	2,404	2,166	MHI-D	2004	60–69	CCSLC		2
de Wit et al. (2009)[a]	.31	Netherlands	1,813	2,020	MHI-D	2004	70–90	CCSLC		2
Dooley et al. (2015)[a]	−.15	Ireland	84	90	DASS-D	2011	12	My World Survey		2
Dooley et al. (2015)[a]	.11	Ireland	462	515	DASS-D	2011	13	My World Survey		2
Dooley et al. (2015)[a]	.22	Ireland	588	609	DASS-D	2011	14	My World Survey		2
Dooley et al. (2015)[a]	.33	Ireland	508	488	DASS-D	2011	15	My World Survey		2
Dooley et al. (2015)[a]	.34	Ireland	451	564	DASS-D	2011	16	My World Survey		2

(*table continues*)

Table 2 (*continued*)

Study	d	Country	NM	NF	Measure	Year	Age	Sample	E	F
Dooley et al. (2015)[a]	.43	Ireland	371	368	DASS-D	2011	17	My World Survey		2
Dooley et al. (2015)[a]	.27	Ireland	142	215	DASS-D	2011	18	My World Survey		2
Everson-Rose et al. (2004)[a]	.31	U.S.	333	407	CES-D	1986	24–34	American Changing Lives Survey	1	2
Everson-Rose et al. (2004)[a]	.23	U.S.	228	363	CES-D	1986	35–44	American Changing Lives Survey	1	2
Everson-Rose et al. (2004)[a]	−.01	U.S.	168	222	CES-D	1986	45–54	American Changing Lives Survey	1	2
Everson-Rose et al. (2004)[a]	.14	U.S.	251	434	CES-D	1986	55–65	American Changing Lives Survey	1	2
Everson-Rose et al. (2004)[a]	.26	U.S.	239	526	CES-D	1986	65–74	American Changing Lives Survey	1	2
Everson-Rose et al. (2004)[a]	.23	U.S.	139	307	CES-D	1986	75+	American Changing Lives Survey	1	2
Ferketich et al. (2000)	.25	U.S.	2,888	5,006	CES-D	1983	30+	National Health & Nutrition Examination Follow-up Study (NHEFS)	1	1
Fleiz Bautista et al. (2012)[a]	.31	Mexico	4,613	4,707	CES-D	2008	12–17	National Survey on Addictions		2
Fleiz Bautista et al. (2012)[a]	.24	Mexico	7,343	7,962	CES-D	2008	18–29	National Survey on Addictions		2
Fleiz Bautista et al. (2012)[a]	.29	Mexico	5,199	5,784	CES-D	2008	30–39	National Survey on Addictions		2
Fleiz Bautista et al. (2012)[a]	.34	Mexico	3,833	4,204	CES-D	2008	40–49	National Survey on Addictions		2
Fleiz Bautista et al. (2012)[a]	.29	Mexico	3,623	3,959	CES-D	2008	50–65	National Survey on Addictions		2
Fleming et al. (2014)[a]	.42	New Zealand	3,074	2,585	RADS	2007	12–15	Youth 2000		2
Fleming et al. (2014)[a]	.32	New Zealand	1,580	1,453	RADS	2007	16–19	Youth 2000		2
Gault-Sherman et al. (2009)[a]	.42	Iceland	1,610	1,596	SCL- D	2004	16	Junior College Questionnaire		1
Gault-Sherman et al. (2009)[a]	.36	Iceland	1,138	1,156	SCL-D	2004	17	Junior College Questionnaire		1
Gault-Sherman et al. (2009)[a]	.38	Iceland	835	927	SCL-D	2004	18	Junior College Questionnaire		1
Gault-Sherman et al. (2009)[a]	.33	Iceland	705	859	SCL-D	2004	19	Junior College Questionnaire		1
Gault-Sherman et al. (2009)[a]	.39	Iceland	314	284	SCL-D	2004	20	Junior College Questionnaire		1
Gettler and Oka (2016)	.14	US	1,505	933	PHQ-9	2012	20–60	NHANES 2011–2012	1	1
Graham et al. (2007)	.15	Canada	2,781	3,675	CES-D	2005	18–76	Gender Alcohol & Culture: An International Study		2
Guarnaccia et al. (1991)	.33	U.S.	1,369	1,583	CES-D	1983	20–45	HHANES	5	2
Guarnaccia et al. (1991)	.31	U.S.	416	659	CES-D	1983	20–45	HHANES	5	2
Guarnaccia et al. (1991)	.42	U.S.	247	312	CES-D	1983	20–45	HHANES	5	2
Gudmundsdottir et al. (2010)	.23	Iceland	967	955	SCL-D	1998	18–75	—		2
Hardie (2014)	.13	U.S.	1,116	1,046	CES-D	1992	27–35	National Longitudinal Survey of Youth (NLSY) '79	3	2
Hardie (2014)	.23	U.S.	725	689	CES-D	1992	27–35	NLSY '79	5	2
Hardie (2014)	.22	U.S.	1,841	1,703	CES-D	1992	27–35	NLSY '79	2	2
Haroz et al. (2014)	.45	U.S.	585	785	CES-D	2014	13–14	Teen Health and Technology	1	2
Haroz et al. (2014)	.27	U.S.	856	1,096	CES-D	2014	15–17	Teen Health and Technology	1	2
Haroz et al. (2014)	.22	U.S.	954	1,404	CES-D	2014	18	Teen Health and Technology	1	2
Haroz et al. (2014)	.16	U.S.	95	99	CES-D	2008	11–12	Growing up with Media	1	2
Haroz et al. (2014)	.47	U.S.	201	191	CES-D	2008	13–14	Growing up with Media	1	2
Haroz et al. (2014)	.11	U.S.	192	172	CES-D	2008	15–17	Growing up with Media	1	2
Haroz et al. (2014)	.25	U.S.	94	106	CES-D	2012	18	Growing up with Media	1	2
Haukkala et al. (2009)[a]	.19	Finland	622	833	BDI	1995	24–34	Finrisk-1995		1
Haukkala et al. (2009)[a]	.15	Finland	783	970	BDI	1995	35–44	Finrisk-1995		1
Haukkala et al. (2009)[a]	.07	Finland	885	991	BDI	1995	45–54	Finrisk-1995		1
Haukkala et al. (2009)[a]	.07	Finland	925	983	BDI	1995	55–64	Finrisk-1995		1
Haukkala et al. (2009)[a]	.22	Finland	493	189	BDI	1995	65–74	Finrisk-1995		1
Hauser et al. (2014)[a]	−.07	Germany	128	144	BDI	2012	14–24	Survey zu psychischen Befindlichkeiten Uni Leipzig		2
Hauser et al. (2014)[a]	.08	Germany	166	181	BDI	2012	25–34	Survey zu psychischen Befindlichkeiten Uni Leipzig		2
Hauser et al. (2014)[a]	.15	Germany	165	195	BDI	2012	35–44	Survey zu psychischen Befindlichkeiten Uni Leipzig		2
Hauser et al. (2014)[a]	−.09	Germany	208	252	BDI	2012	45–54	Survey zu psychischen Befindlichkeiten Uni Leipzig		2
Hauser et al. (2014)[a]	.07	Germany	212	250	BDI	2012	55–64	Survey zu psychischen Befindlichkeiten Uni Leipzig		2
Hauser et al. (2014)[a]	−.07	Germany	209	195	BDI	2012	65–74	Survey zu psychischen Befindlichkeiten Uni Leipzig		2
Hauser et al. (2014)[a]	.18	Germany	76	113	BDI	2012	85+	Survey zu psychischen Befindlichkeiten Uni Leipzig		2
Huang and Chen (2015)[a]	.45	Taiwan	52	64	CES-D	2010	15	—		2
Huang and Chen (2015)[a]	−.01	Taiwan	253	283	CES-D	2010	16	—		2
Huang and Chen (2015)[a]	.10	Taiwan	205	206	CES-D	2010	17	—		2
Huang and Chen (2015)[a]	.32	Taiwan	43	90	CES-D	2010	18	—		2
Hwang et al. (2009)[a]	.17	Taiwan	2,904	2,966	—	2004	12–17	Adolescent Internet Use, Daily Life, & Depressive Mood Survey		2
Inaba et al. (2005)[a]	.09	Japan	740	834	CES-D	1999	28–39	National Family Research of Japan '98 Survey (NFRJ98)		1
Inaba et al. (2005)[a]	.15	Japan	675	732	CES-D	1999	40–49	NFRJ98		1
Inaba et al. (2005)[a]	.18	Japan	670	794	CES-D	1999	50–59	NFRJ98		1
Inaba et al. (2005)[a]	.17	Japan	906	977	CES-D	1999	60–78	NFRJ98		1
Inaba et al. (2005)[a]	.27	U.S.	1,372	1,413	CES-D	1994	28–39	National Survey of Families & Households (NSFH) in the US	1	1
Inaba et al. (2005)[a]	.22	U.S.	1,013	987	CES-D	1994	40–49	NSFH	1	1
Inaba et al. (2005)[a]	.38	U.S.	594	716	CES-D	1994	50–59	NSFH	1	1
Inaba et al. (2005)[a]	.29	U.S.	856	1,220	CES-D	1994	60–78	NSFH	1	1

Table 2 (*continued*)

Study	d	Country	NM	NF	Measure	Year	Age	Sample	E	F
Ismayilova et al. (2013)[a]	.32	Jordan	3,513	3,252	—	2009	14–18	—		2
Jaddou et al. (2012)[a]	.36	Jordan	1,008	2,994	DASS-D	2009	25+	—		2
Jang et al. (2009)	.14	South Korea	1,320	1,648	CES-D	2006	45–54	Korean Longitudinal Study of Aging		1
Jang et al. (2009)	.19	South Korea	1,163	1,341	CES-D	2006	55–64	Korean Longitudinal Study of Aging		1
Jang et al. (2009)	.32	South Korea	1,095	1,306	CES-D	2006	65–74	Korean Longitudinal Study of Aging		1
Jang et al. (2009)	.26	South Korea	483	708	CES-D	2006	75–85	Korean Longitudinal Study of Aging		1
Kaji et al. (2010)[a]	.10	Japan	2,543	2,583	CES-D	2000	50–59	Active Survey of Health & Welfare in Japan		2
Kaji et al. (2010)[a]	.04	Japan	1,691	1,745	CES-D	2000	60–69	Active Survey of Health & Welfare in Japan		2
Kaji et al. (2010)[a]	.13	Japan	799	1,003	CES-D	2000	70–79	Active Survey of Health & Welfare in Japan		2
Kaji et al. (2010)[a]	.10	Japan	200	495	CES-D	2000	80+	Active Survey of Health & Welfare in Japan		2
Karadag et al. (2014)[a]	.34	Turkey	400	342	BDI	2012	18–24	Prevalence of Sleep Disorders in the Turkish Adult Population Epidemiology of Sleep (TAPES) Study		1
Karadag et al. (2014)[a]	.34	Turkey	538	742	BDI	2012	25–34	TAPES Study		1
Karadag et al. (2014)[a]	.26	Turkey	448	661	BDI	2012	35–44	TAPES Study		1
Karadag et al. (2014)[a]	.20	Turkey	451	429	BDI	2012	45–54	TAPES Study		1
Karadag et al. (2014)[a]	.29	Turkey	319	270	BDI	2012	55–64	TAPES Study		1
Karadag et al. (2014)[a]	.44	Turkey	266	152	BDI	2012	65+	TAPES Study		1
Kim (2016)	.30	South Korea	1,137	1,256	—	2003	13–14	Korea Youth Panel Survey		3
Kliem et al. (2014)[a]	.03	Germany	2,106	2,374	BDI	2012	14–91	—		2
Klomek et al. (2009)[a]	−.16	Finland	2,651	2,583	CDI	1989	8	—		2
Kocalevent et al. (2013)	.12	Germany	292	272	PHQ-9	2006	14–24	—		2
Kocalevent et al. (2013)	.16	Germany	279	351	PHQ-9	2006	25–34	—		2
Kocalevent et al. (2013)	.16	Germany	396	542	PHQ-9	2006	35–44	—		2
Kocalevent et al. (2013)	−.05	Germany	414	457	PHQ-9	2006	45–54	—		2
Kocalevent et al. (2013)	.04	Germany	398	446	PHQ-9	2006	55–64	—		2
Kocalevent et al. (2013)	.18	Germany	397	395	PHQ-9	2006	65–74	—		2
Kocalevent et al. (2013)	.11	Germany	156	236	PHQ-9	2006	75+	—		2
Kokkevi & Fotiou (2009)[a]	.43	Armenia	1,569	2,231	CES-D	2007	16	European School Survey Project on Alcohol & Other Drugs (ESPAD)		3
Kokkevi & Fotiou (2009)[a]	.59	Bulgaria	1,098	1,056	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.66	Croatia	1,497	1,438	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.53	Cyprus	2,936	3,182	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.42	Faroe Islands	247	281	CESD	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.69	Finland	2,242	2,641	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.54	Hungary	1,277	1,386	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.48	Iceland	1,634	1,629	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.59	Ireland	932	1,136	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.50	Isle of Man	356	362	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.57	Latvia	1,059	1,135	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.70	Romania	943	1,231	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.60	Slovak Rep.	1,166	1,211	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.70	Slovenia	1,525	1,466	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.56	U.K.	909	1,120	CES-D	2007	16	ESPAD		3
Kokkevi & Fotiou (2009)[a]	.52	Greece	4,631	5,242	CES-D	2007	14–18	ESPAD		2
Kokkevi et al. (2011)[a]	.61	Greece	5,169	5,336	CES-D	1984	14–18	Greek National School Population Survey		2
Kopp et al. (1995)	.07	Hungary	7,767	7,979	BDI	1988	16–65	—		2
Lee (2015)[a]	.16	South Korea	617	779	PHQ-9	2012	50.6	Korean General Social Survey		2
Lei et al. (2014)	.20	China	1,285	1,740	CES-D	2012	45–49	China Health & Retirement Longitudinal Study (CHARLS)		2
Lei et al. (2014)	.24	China	1,050	1,182	CES-D	2012	50–54	CHARLS		2
Lei et al. (2014)	.26	China	1,484	1,591	CES-D	2012	55–59	CHARLS		2
Lei et al. (2014)	.31	China	1,294	1,270	CES-D	2012	60–64	CHARLS		2
Lei et al. (2014)	.30	China	839	801	CES-D	2012	65–69	CHARLS		2
Lei et al. (2014)	.22	China	639	569	CES-D	2012	70–74	CHARLS		2
Lei et al. (2014)	.37	China	585	596	CES-D	2012	75+	CHARLS		2
Madianos et al. (1992)	.45	Greece	1,820	2,263	CES-D	1978	19–64	—		2
Madianos et al. (1992)	.33	Greece	1,660	2,046	CES-D	1984	19–64	—		2
Mäntyselkä et al. (2003)[a]	.35	Finland	258	389	DEPS	2002	15–24	Finnish Pain Study		3
Mäntyselkä et al. (2003)[a]	.17	Finland	278	378	DEPS	2002	25–34	Finnish Pain Study		3
Mäntyselkä et al. (2003)[a]	.12	Finland	309	399	DEPS	2002	35–44	Finnish Pain Study		3
Mäntyselkä et al. (2003)[a]	−.08	Finland	336	432	DEPS	2002	45–54	Finnish Pain Study		3
Mäntyselkä et al. (2003)[a]	.16	Finland	381	441	DEPS	2002	55–64	Finnish Pain Study		3
Mäntyselkä et al. (2003)[a]	.14	Finland	417	437	DEPS	2002	65–74	Finnish Pain Study		3
Margraf et al. (2016)[a]	−.12	U.S.	1,252	1,786	DASS-D	2013	18+	Bochum Optimism and Mental Health (BOOM)- US	1	3
Margraf et al. (2016)[a]	.08	Germany	826	1,181	DASS-D	2013	18+	BOOM- Germany		3
Margraf et al. (2016)[a]	.13	Russia	1,413	1,607	DASS-D	2013	18+	BOOM- Russia		3
Marmorstein (2009)[a]	.12	U.S.	262	329	CES-D	1995	12	National Longitudinal Student of Adolescent Health (Add Health)	1	2
Marmorstein (2009)[a]	.22	U.S.	1,039	1,218	CES-D	1995	13	Add Health	1	2

(*table continues*)

Table 2 (*continued*)

Study	*d*	Country	NM	NF	Measure	Year	Age	Sample	E	F
Marmorstein (2009)[a]	.30	U.S.	1,319	1,472	CES-D	1995	14	Add Health	1	2
Marmorstein (2009)[a]	.34	U.S.	1,778	1,883	CES-D	1995	15	Add Health	1	2
Marmorstein (2009)[a]	.31	U.S.	2,061	1,991	CES-D	1995	16	Add Health	1	2
Marmorstein (2009)[a]	.19	U.S.	1,981	1,940	CES-D	1995	17	Add Health	1	2
Marmorstein (2009)[a]	.21	U.S.	1,512	1,427	CES-D	1995	18	Add Health	1	2
Marmorstein (2009)[a]	.34	U.S.	237	159	CES-D	1995	19	Add Health	1	2
Maske et al. (2016)[a]	.27	Germany	722	781	PHQ-9	2010	18–34	Germany Health Interview and Examination Survey for Adults (DEGS1)		2
Maske et al. (2016)[a]	.28	Germany	553	646	PHQ-9	2010	35–44	DEGS1		2
Maske et al. (2016)[a]	.28	Germany	741	858	PHQ-9	2010	45–54	DEGS1		2
Maske et al. (2016)[a]	.25	Germany	652	718	PHQ-9	2010	55–64	DEGS1		2
Maske et al. (2016)[a]	.30	Germany	915	937	PHQ-9	2010	65–79	DEGS1		2
Michal et al. (2011)[a]	−.07	Germany	1,090	1,385	HADS-D	2009	14–94	—		2
Momtaz et al. (2016)[a]	.15	Malaysia	564	696	GDS-15	2014	60–69	Identifying Psychosocial and Identifying Economic Risk Factor of Cognitive Impairment among Elderly		3
Momtaz et al. (2016)[a]	.14	Malaysia	423	387	GDS-15	2014	70–79	Identifying Psychosocial and Identifying Economic Risk Factor of Cognitive Impairment among Elderly		3
Momtaz et al. (2016)[a]	.15	Malaysia	67	65	GDS-15	2014	80+	Identifying Psychosocial and Identifying Economic Risk Factor of Cognitive Impairment among Elderly		3
Morozink et al. (2010)[a]	.10	U.S.	541	705	CES-D	2004	35–86	Midlife in the United States (MIDUS)	1	2
Mumford et al. (2013)[a]	.48	U.S.	815	765	MHI-D	2000	15	National Longitudinal Survey of Youth (NLSY) '97	1	2
Mumford et al. (2013)[a]	.33	U.S.	819	774	MHI-D	2000	16	NLSY '97	1	2
Mumford et al. (2013)[a]	.29	U.S.	811	773	MHI-D	2000	17	NLSY '97	1	2
Mumford et al. (2013)[a]	.22	U.S.	766	767	MHI-D	2000	18	NLSY '97	1	2
Mumford et al. (2013)[a]	.23	U.S.	657	681	MHI-D	2000	19	NLSY '97	1	2
Munhoz et al. (2016)[a]	.32	Brazil	6,246	8,075	PHQ-9	2013	18–29	Brazilian National Healthy Survey		2
Munhoz et al. (2016)[a]	.34	Brazil	6,049	8,220	PHQ-9	2013	30–39	Brazilian National Healthy Survey		2
Munhoz et al. (2016)[a]	.37	Brazil	5,079	6,326	PHQ-9	2013	40–49	Brazilian National Healthy Survey		2
Munhoz et al. (2016)[a]	.37	Brazil	3,991	5,039	PHQ-9	2013	50–59	Brazilian National Healthy Survey		2
Munhoz et al. (2016)[a]	.26	Brazil	2,595	3,643	PHQ-9	2013	60–69	Brazilian National Healthy Survey		2
Munhoz et al. (2016)[a]	.27	Brazil	1,428	2,013	PHQ-9	2013	70–79	Brazilian National Healthy Survey		2
Munhoz et al. (2016)[a]	.14	Brazil	532	966	PHQ-9	2013	80+	Brazilian National Healthy Survey		2
Neumark-Sztainer et al. (2000)[a]	.12	U.S.	239	267	CDI	1997	10	Commonwealth Fund Survey (of the Health of Adolescent Girls & Boys)	1	1
Neumark-Sztainer et al. (2000)[a]	.02	U.S.	254	305	CDI	1997	11	Commonwealth Fund Survey	1	1
Neumark-Sztainer et al. (2000)[a]	.04	U.S.	386	461	CDI	1997	12	Commonwealth Fund Survey	1	1
Neumark-Sztainer et al. (2000)[a]	.29	U.S.	420	484	CDI	1997	13	Commonwealth Fund Survey	1	1
Neumark-Sztainer et al. (2000)[a]	.22	U.S.	370	462	CDI	1997	14	Commonwealth Fund Survey	1	1
Neumark-Sztainer et al. (2000)[a]	.31	U.S.	361	503	CDI	1997	15	Commonwealth Fund Survey	1	1
Neumark-Sztainer et al. (2000)[a]	.32	U.S.	399	497	CDI	1997	16	Commonwealth Fund Survey	1	1
Neumark-Sztainer et al. (2000)[a]	.25	U.S.	314	372	CDI	1997	17	Commonwealth Fund Survey	1	1
O'Halloran et al. (2014)[a]	.21	Ireland	603	727	CES-D	2010	50–54	The Irish Longitudinal Study on Ageing (TILDA)		2
O'Halloran et al. (2014)[a]	.18	Ireland	591	754	CES-D	2010	55–59	TILDA		2
O'Halloran et al. (2014)[a]	.20	Ireland	502	647	CES-D	2010	60–64	TILDA		2
O'Halloran et al. (2014)[a]	.19	Ireland	487	502	CES-D	2010	65–69	TILDA		2
O'Halloran et al. (2014)[a]	.35	Ireland	383	376	CES-D	2010	70–74	TILDA		2
O'Halloran et al. (2014)[a]	.27	Ireland	449	521	CES-D	2010	75+	TILDA		2
Oh et al. (2013)[a]	.30	South Korea	14,452	16,146	CES-D	2009	19–29	Korean Community Health Survey		2
Oh et al. (2013)[a]	.19	South Korea	20,031	21,340	CES-D	2009	30–39	Korean Community Health Survey		2
Oh et al. (2013)[a]	.17	South Korea	23,381	24,288	CES-D	2009	40–49	Korean Community Health Survey		2
Oh et al. (2013)[a]	.22	South Korea	19,781	21,722	CES-D	2009	50–59	Korean Community Health Survey		2
Oh et al. (2013)[a]	.39	South Korea	15,986	19,191	CES-D	2009	60–69	Korean Community Health Survey		2
Oh et al. (2013)[a]	.35	South Korea	12,912	20,365	CES-D	2009	70+	Korean Community Health Survey		2
Ojard et al. (2015)[a]	.29	U.S.	8,751	8,802	CES-D	2005	45+	Reasons for Geographic & Racial Differences in Stroke (REGARDS)	2	2
Ojard et al. (2015)[a]	.19	U.S.	4,705	7,710	CES-D	2005	45+	REGARDS	3	2
Okabayashi et al. (2004)[a]	.21	Japan	995	1,205	CES-D	1987	60+	—		3
Oksuzyan et al. (2010)[a]	.13	U.S.	640	1,013	CES-D	2006	50–54	US Health & Retirement Survey	1	1
Oksuzyan et al. (2010)[a]	.05	U.S.	1,051	1,472	CES-D	2006	55–59	US Health & Retirement Survey	1	1
Oksuzyan et al. (2010)[a]	.10	U.S.	936	1,463	CES-D	2006	60–64	US Health & Retirement Survey	1	1
Oksuzyan et al. (2010)[a]	.11	U.S.	1,537	1,879	CES-D	2006	65–69	US Health & Retirement Survey	1	1
Oksuzyan et al. (2010)[a]	.12	U.S.	1,267	1,560	CES-D	2006	70–74	US Health & Retirement Survey	1	1

Table 2 (*continued*)

Study	d	Country	NM	NF	Measure	Year	Age	Sample	E	F
Oksuzyan et al. (2010)[a]	.16	U.S.	906	1,128	CES-D	2006	75–79	US Health & Retirement Survey	1	1
Oksuzyan et al. (2010)[a]	.11	U.S.	647	917	CES-D	2006	80–84	US Health & Retirement Survey	1	1
Oksuzyan et al. (2010)[a]	.04	U.S.	344	649	CES-D	2006	85–89	US Health & Retirement Survey	1	1
Oksuzyan et al. (2010)[a]	.04	U.S.	142	379	CES-D	2006	90+	US Health & Retirement Survey	1	1
Oksuzyan et al. (2010)[a]	.01	Japan	181	213	CES-D	2006	67–69	Nihon University Japanese Longitudinal Study of Aging (NUJLSOA)		1
Oksuzyan et al. (2010)[a]	.00	Japan	435	469	CES-D	2006	70–74	NUJLSOA		1
Oksuzyan et al. (2010)[a]	−.03	Japan	302	329	CES-D	2006	75–79	NUJLSOA		1
Oksuzyan et al. (2010)[a]	−.03	Japan	263	373	CES-D	2006	80–84	NUJLSOA		1
Oksuzyan et al. (2010)[a]	.04	Japan	133	217	CES-D	2006	85–89	NUJLSOA		1
Olsen et al. (2006)	.31	Denmark	121	165	SCL-D	2000	19–34	—		2
Olsen et al. (2006)	.29	Denmark	133	184	SCL-D	2000	35–49	—		2
Olsen et al. (2006)	.28	Denmark	135	150	SCL-D	2000	50–64	—		2
Olsen et al. (2006)	.09	Denmark	143	121	SCL-D	2000	65–80	—		2
Ostbye et al. (2009)[a]	.20	Sri Lanka	463	623	GDS	2006	60+	National Sri Lanka Aging Survey		2
Park et al. (2012)[a]	−.04	South Korea	98	309	GDS	2008	85+	Nationwide Survey on Dementia in Korea		2
Park et al. (2012)[a]	.13	South Korea	180	447	GDS	2008	80–84	Nationwide Survey on Dementia in Korea		2
Park et al. (2012)[a]	.17	South Korea	455	740	GDS	2008	75–79	Nationwide Survey on Dementia in Korea		2
Park et al. (2012)[a]	.36	South Korea	790	1,037	GDS	2008	70–74	Nationwide Survey on Dementia in Korea		2
Park et al. (2012)[a]	.33	South Korea	885	1,077	GDS	2008	64–69	Nationwide Survey on Dementia in Korea		2
Revah-Levy et al. (2011)	.43	France	19,884	19,658	ADRS	2008	17	ESCAPAD 2008 (Enquête sur lar Santeʹ et les Consommations lors de lʹAppel de Preʹparation aʹ la Deʹfense)		1
Rey et al. (2001)[a]	.30	Australia	594	654	CES-D	1998	13–17	National Survey of Mental Health & Well-being		2
Rief et al. (2012)[a]	.06	Germany	1,202	1,316	PHQ-9	2008	14+	—		2
Risal et al. (2016)[a]	.02	Nepal	189	300	HADS-D	2013	18–25	—		2
Risal et al. (2016)[a]	.32	Nepal	217	440	HADS-D	2013	26–35	—		2
Risal et al. (2016)[a]	.20	Nepal	196	242	HADS-D	2013	36–45	—		2
Risal et al. (2016)[a]	.37	Nepal	138	160	HADS-D	2013	46–55	—		2
Risal et al. (2016)[a]	.09	Nepal	121	97	HADS-D	2013	56–65	—		2
Sandman et al. (2015)[a]	.25	Finland	263	361	BDI	2007	25–34	Finrisk-2007		3
Sandman et al. (2015)[a]	.26	Finland	318	392	BDI	2007	35–44	Finrisk-2007		3
Sandman et al. (2015)[a]	.14	Finland	371	442	BDI	2007	45–54	Finrisk-2007		3
Sandman et al. (2015)[a]	.07	Finland	412	452	BDI	2007	55–64	Finrisk-2007		3
Sandman et al. (2015)[a]	.16	Finland	407	420	BDI	2007	65–74	Finrisk-2007		3
Sandman et al. (2015)[a]	.29	Finland	260	410	BDI	2012	25–34	Finrisk-2012		3
Sandman et al. (2015)[a]	.30	Finland	338	469	BDI	2012	35–44	Finrisk-2012		3
Sandman et al. (2015)[a]	.11	Finland	430	534	BDI	2012	45–54	Finrisk-2012		3
Sandman et al. (2015)[a]	.20	Finland	503	580	BDI	2012	55–64	Finrisk-2012		3
Sandman et al. (2015)[a]	.16	Finland	613	592	BDI	2012	65–74	Finrisk-2012		3
Scafato et al. (2012)[a]	.56	Italy	1,713	1,501	GDS	1993	65–84	Italian Longitudinal Study on Aging		2
Seaton et al. (2008)	.02	U.S.	563	605	CES-D	2002	13–17	National Survey of African Life	3	2
Seppälä et al. (2012)[a]	.19	Finland	391	483	BDI	2007	45–54	FIN-D2D		2
Seppälä et al. (2012)[a]	.21	Finland	450	511	BDI	2007	55–64	FIN-D2D		2
Seppälä et al. (2012)[a]	.23	Finland	487	484	BDI	2007	65–74	FIN-D2D		2
Shiovitz-Ezra et al. (2009)[a]	.23	U.S.	521	484	CES-D	2006	57–64	National Social Life, Health, & Aging Project (NSHAP)	1	2
Shiovitz-Ezra et al. (2009)[a]	.16	U.S.	543	537	CES-D	2006	65–74	NSHAP	1	2
Shiovitz-Ezra et al. (2009)[a]	.09	U.S.	373	499	CES-D	2006	75–85	NSHAP	1	2
Sigfusdottir et al. (2008)	.49	Iceland	1,927	1,802	SCL-D	1997	14–15	Youth in Iceland-1997		2
Sigfusdottir et al. (2008)	.39	Iceland	2,908	3,132	SCL-D	2000	14–15	Youth in Iceland-2000		2
Sigfusdottir et al. (2008)	.45	Iceland	1,709	1,662	SCL-D	2003	14–15	Youth in Iceland-2003		2
Sigfusdottir et al. (2008)	.55	Iceland	3,437	3,503	SCL-D	2006	14–15	Youth in Iceland-2006		2
Song (2011)[a]	.11	U.S.	939	1,124	CES-D	2005	21–64	—	2	3
Song (2011)[a]	.30	U.S.	167	188	CES-D	2005	21–64	—	3	3
Song (2011)[a]	.21	U.S.	187	225	CES-D	2005	21–64	—	5	3
Sonnenberg et al. (2013)[a]	.26	Netherlands	449	483	CES-D	1993	55–64	Longitudinal Aging Study Amsterdam		1
Sonnenberg et al. (2013)[a]	.44	Netherlands	423	476	CES-D	1993	65–74	Longitudinal Aging Study Amsterdam		1
Sonnenberg et al. (2013)[a]	.31	Netherlands	497	495	CES-D	1993	75–85	Longitudinal Aging Study Amsterdam		1
Steptoe et al. (2012)	.28	England	3,540	4,255	CES-D	2009	50–90	English Longitudinal Study of Ageing		1
Strong et al. (2016)[a]	.22	Taiwan	2,113	2,148	—	2001	13	Taiwan Education Panel Survey		2
Sugihara et al. (2008)	.12	Japan	2,533	1,440	CES-D	1999	55–64	Japanese Health & Retirement Study		1
Symonds et al. (2016)[a]	.42	England	6,549	6,422	GHQ-D	2005	15	Longitudinal Study of Young People in England		3
Thege et al. (2009)[a]	.14	Hungary	5,416	6,709	BDI	2002	18–98	Hungaro Study		3
Thibodeau et al. (2014)[a]	.29	U.S.	550	500	PHQ-9	2008	18–29	National Health & Nutrition Examination Survey (NHANES)- 2008	1	1
Thibodeau et al. (2014)[a]	.34	U.S.	431	447	PHQ-9	2008	30–39	NHANES-2008	1	1
Thibodeau et al. (2014)[a]	.30	U.S.	391	452	PHQ-9	2008	40–49	NHANES-2008	1	1
Thibodeau et al. (2014)[a]	.23	U.S.	418	400	PHQ-9	2008	50–59	NHANES-2008	1	1
Thibodeau et al. (2014)[a]	.29	U.S.	434	459	PHQ-9	2008	60–69	NHANES-2008	1	1
Thibodeau et al. (2014)[a]	.25	U.S.	483	482	PHQ-9	2008	70+	NHANES-2008	1	1
Torikka et al. (2014)[a]	.35	Finland	47,586	47,049	BDI	2001	14–16	School Health Promotion Study, 2000–2001		2

(*table continues*)

Table 2 (*continued*)

Study	d	Country	NM	NF	Measure	Year	Age	Sample	E	F
Torikka et al. (2014)[a]	.34	Finland	50,744	49,236	BDI	2003	14–16	School Health Promotion Study (SHPS), 2002–2003		2
Torikka et al. (2014)[a]	.35	Finland	53,057	51,713	BDI	2005	14–16	SHPS 2004–2005		2
Torikka et al. (2014)[a]	.33	Finland	54,315	54,005	BDI	2007	14–16	SHPS, 2006–2007		2
Torikka et al. (2014)[a]	.33	Finland	54,132	54,035	BDI	2009	14–16	SHPS, 2008–2009		2
Torikka et al. (2014)[a]	.37	Finland	51,116	51,066	BDI	2011	14–16	SHPS, 2010–2011		2
Torres and Wong (2013)[a]	.38	Mexico	1,094	1,523	CES-D	2001	60–69	Mexican Health & Aging Study		2
Torres and Wong (2013)[a]	.38	Mexico	672	843	CES-D	2001	70–79	Mexican Health & Aging Study		2
Torres and Wong (2013)[a]	.27	Mexico	208	330	CES-D	2001	80+	Mexican Health & Aging Study		2
Torres and Wong (2013)[a]	.49	Mexico	2,027	2,166	CES-D	2001	50–59	Mexican Health & Aging Study		2
Van de Velde et al. (2010)[a]	.14	Austria	1,099	1,285	CES-D	2007	15–99	European Social Survey (ESS)- 3		1
Van de Velde et al. (2010)[a]	.33	Belgium	839	958	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.21	Bulgaria	544	828	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.17	Switzerland	840	963	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.42	Cyprus	471	523	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.21	Germany	1,437	1,473	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.16	Denmark	729	756	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.21	Estonia	658	855	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.35	Spain	905	969	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.09	Finland	915	976	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.29	France	968	1018	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.17	U.K.	1,137	1,257	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.18	Hungary	640	874	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.03	Ireland	812	928	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.31	Netherlands	897	991	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.15	Norway	889	859	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.29	Poland	815	896	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.37	Portugal	911	1,309	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.35	Russian Fed	994	1,395	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.28	Sweden	948	973	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.23	Slovenia	665	807	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.13	Slovakia	839	895	CES-D	2007	15–99	ESS-3		1
Van de Velde et al. (2010)[a]	.21	Ukraine	849	1,136	CES-D	2007	15–99	ESS-3		1
Van de Velde (personal communication, April 2, 2015)[a]	.27	Belgium	910	958	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.24	Bulgaria	971	1,282	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.41	Cyprus	485	631	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.23	Czech Republic	1,012	977	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.20	Denmark	832	814	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.13	Estonia	999	1,379	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.09	Finland	1,074	1,121	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.24	Germany	1,487	1,469	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.06	Ireland	1,266	1,357	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.09	Israel	1,142	1,354	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.25	Kosovo	619	676	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.24	Netherlands	866	979	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.14	Norway	855	763	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.31	Poland	908	985	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.28	Portugal	867	1,284	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.27	Russian Fed.	978	1,484	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.08	Slovakia	795	1,046	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.22	Slovenia	572	684	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.35	Spain	912	975	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.20	Sweden	947	900	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.28	Switzerland	746	747	CES-D	2013	15–99	ESS-6		2
Van de Velde (2015)[a]	.15	UK	990	1,293	CES-D	2013	15–99	ESS-6		2
van Praag et al. (2009)	.29	Belgium	9,378	10,085	SCL-D	2003	15+	Belgian Health Interview Survey		1
Villatoro et al. (1998)[a]	.08	Mexico	5,319	5,511	CES-D	1991	12	1991 National School Survey		2
Villatoro et al. (1998)[a]	.24	Mexico	7,002	6,819	CES-D	1991	13	1991 National School Survey		2
Villatoro et al. (1998)[a]	.29	Mexico	6,305	5,768	CES-D	1991	14	1991 National School Survey		2
Villatoro et al. (1998)[a]	.35	Mexico	4,848	4,000	CES-D	1991	15	1991 National School Survey		2
Villatoro et al. (1998)[a]	.36	Mexico	3,349	2,971	CES-D	1991	16	1991 National School Survey		2
Villatoro et al. (1998)[a]	.42	Mexico	2,197	1,725	CES-D	1991	17	1991 National School Survey		2
Villatoro et al. (1998)[a]	.39	Mexico	991	635	CES-D	1991	18	1991 National School Survey		2
Villatoro et al. (1998)[a]	.28	Mexico	431	297	CES-D	1991	19	1991 National School Survey		2
von Soest & Wichstrøm (2014)[a]	.21	Norway	724	802	SCL-D	2002	13	Young in Norway		2
von Soest et al. (2014)[a]	.32	Norway	899	950	SCL-D	2002	14	Young in Norway		2
von Soest et al. (2014)[a]	.38	Norway	932	959	SCL-D	2002	15	Young in Norway		2
von Soest et al. (2014)[a]	.41	Norway	704	698	SCL-D	2002	16	Young in Norway		2
von Soest et al. (2014)[a]	.42	Norway	1,004	992	SCL-D	2002	17	Young in Norway		2
von Soest et al. (2014)[a]	.42	Norway	714	906	SCL-D	2002	18	Young in Norway		2
von Soest et al. (2014)[a]	.21	Norway	418	416	SCL-D	2010	12	Young in Norway		2
von Soest et al. (2014)[a]	.21	Norway	694	761	SCL-D	2010	13	Young in Norway		2

Table 2 (*continued*)

Study	*d*	Country	NM	NF	Measure	Year	Age	Sample	E	F
von Soest et al. (2014)[a]	.38	Norway	610	653	SCL-D	2010	14	Young in Norway		2
von Soest et al. (2014)[a]	.52	Norway	573	695	SCL-D	2010	15	Young in Norway		2
von Soest et al. (2014)[a]	.45	Norway	621	612	SCL-D	2010	16	Young in Norway		2
von Soest et al. (2014)[a]	.41	Norway	76	714	SCL-D	2010	17	Young in Norway		2
Walker et al. (2005)	.13	New Zealand	4,266	5,049	RADS	2001	12–18	Youth 2000		2
Wang et al. (2010)[a]	.29	U.S.	1,164	1,186	—	2006	11	Health Behavior in School-Aged Children (HBSC)	1	2
Wang et al. (2010)[a]	.44	U.S.	892	951	—	2006	12	HBSC	1	2
Wang et al. (2010)[a]	.49	U.S.	789	997	—	2006	13	HBSC	1	2
Wang et al. (2010)[a]	.51	U.S.	721	742	—	2006	14	HBSC	1	2
Wang et al. (2010)[a]	.59	U.S.	793	804	—	2006	15	HBSC	1	2
Wichstrøm (1999)[a]	−.09	Norway	37	64	SCL-D	1992	12	Young in Norway		1
Wichstrøm (1999)[a]	.18	Norway	628	689	SCL-D	1992	13	Young in Norway		1
Wichstrøm (1999)[a]	.44	Norway	827	802	SCL-D	1992	14	Young in Norway		1
Wichstrøm (1999)[a]	.50	Norway	870	869	SCL-D	1992	15	Young in Norway		1
Wichstrøm (1999)[a]	.58	Norway	919	898	SCL-D	1992	16	Young in Norway		1
Wichstrøm (1999)[a]	.54	Norway	754	725	SCL-D	1992	17	Young in Norway		1
Wichstrøm (1999)[a]	.58	Norway	665	710	SCL-D	1992	18	Young in Norway		1
Wichstrøm (1999)[a]	.45	Norway	271	300	SCL-D	1992	19	Young in Norway		1
Wichstrøm (1999)[a]	.59	Norway	96	161	SCL-D	1992	20	Young in Norway		1
Yamada and Teerawichitchainan (2015)[a]	.37	Vietnam	458	598	CES-D	2011	60–69	Vietnam Aging Survey-2011		3
Yamada et al. (2015)[a]	.53	Vietnam	280	395	CES-D	2011	70–79	Vietnam Aging Survey-2011		3
Yamada et al. (2015)[a]	.39	Vietnam	194	300	CES-D	2011	80+	Vietnam Aging Survey-2011		3
Zemore et al. (2013)[a]	.13	U.S.	2306	2599	CES-D	2000	18+	National Alcohol Surveys (NAS)	2	3
Zemore et al. (2013)[a]	.13	U.S.	514	847	CES-D	2000	18+	NAS	3	3
Zemore et al. (2013)[a]	.11	U.S.	464	530	CES-D	2000	18+	NAS	5	3
Zemore et al. (2013)[a]	.16	U.S.	1,903	2,064	CES-D	2005	18+	NAS	2	3
Zemore et al. (2013)[a]	.09	U.S.	383	671	CES-D	2005	18+	NAS	3	3
Zemore et al. (2013)[a]	.09	U.S.	784	826	CES-D	2005	18+	NAS	5	3
Zemore et al. (2013)[a]	.15	U.S.	1,904	2,695	CES-D	2010	18+	NAS	2	3
Zemore et al. (2013)[a]	.10	U.S.	517	1,078	CES-D	2010	18+	NAS	3	3
Zemore et al. (2013)[a]	.09	U.S.	517	936	CES-D	2010	18+	NAS	5	3
Zunzunegui et al. (2007)	.52	Israel	523	550	CES-D	1989	75–84	Cross Sectional & Longitudinal Aging Study		1

Note. d = uncorrected effect size; NM = n men; NF = n women; Measure = depression symptom measure (the version of each measure- including number of items, language, and response scale- varied for each sample); Year = year of data collection; Age = age or age range (in years); Sample = sample name (not all samples had a name); E = ethnicity (only applies to U.S. samples): 1 = mixed, 2 ≥ 85% White, 3 ≥ 85% African American, 4 ≥ 85% Asian American, 5 ≥ 85% Hispanic; F = focus of article: 1 = gender, 2 = depression, 3 = other; [a] = received additional data from author; BDI = Beck Depression Inventory (Beck et al., 1961, 1996); DMI = Depressive Mood Inventory (see Abebe et al., 2016); CDI = Children's Depression Inventory (Kovacs, 1985); MDI = Major Depression Inventory (Bech et al., 2001); DTS = Depressive Tendencies Scale (Alsaker et al., 1991); HDL-D = Health and Daily Living Form- Depression Scale (Moos et al., 1985); CES-D = Center for Epidemiological Depression Scale (Radloff, 1977); SCL-D = Symptom Checklist-90- Depression items (Derogatis et al., 1973); HADS-D = Hospital Anxiety and Depression Scale- Depression subscale (Zigmond & Snaith, 1983); RADS = Reynolds Adolescent Depression Scale (Reynolds, 1986); EURO-D (Prince et al., 1999); MHI-D = Mental Health Inventory- Depression Items (Ware & Sherbourne, 1992); DASS-D = Depression Anxiety Stress Scales- Depression Subscale (Lovibond & Lovibond, 1995); PHQ-9 = Patient Health Questionnaire- 9 (Kroenke et al., 2001); DEPS = Depression Scale (Salokangas et al., 1995); GDS = Geriatric Depression Scale (Yesavage et al., 1982); ADRS = Adolescent Depression Rating Scale (Revah-Levy et al., 2007); GHQ-D = General Health Questionnaire-Depression items (Symonds et al., 2016).

data are not available for all nations. Therefore, sample sizes vary in regression analyses, depending on the indicator.

Nation-Level Gender Equity Indicators

We selected the following five domain-specific nation-level gender equity indicators as hypothesized moderators: contraceptive prevalence (percentage of women in a nation aged 15–49 using some form of contraception, an indicator of women's ability to control their reproduction), executive positions (percentage of executive positions held by women, a measure of women's equality in the workplace), literacy ratio (female:male ratio in percentage of the adult population that is literate, a measure of women's equality in education), intimate-partner violence against women (lifetime prevalence of physical violence against women by intimate partners), and sexism ideals (measured in response to the item "When jobs are scarce, men should have more right to a job than women").

Contraceptive prevalence, executive positions, and literacy indicators for each country were obtained from the Human Development Report, 2005 (United Nations Development Programme, 2005). The intimate-partner violence measure came from the United Nations' 2010 report, *The World's Women*, for the years 2000–2006. The sexism ideals measure was obtained from the World Values Survey, Wave 4, 1999–2004 (World Values Survey, 2014). Although they were theoretically interesting, the intimate-partner violence and sexism ideals measures proved unsatisfactory because they were available for only a minority of the effect sizes in analyses. Those two measures are therefore not considered further. See supplemental material Table 1 for a list of each country and its associated indicators.

Effect Size Computation

The *OR* (Lipsey & Wilson, 2001) was the effect size used for the meta-analysis on major depression diagnoses; Cohen's *d* (Cohen,

1988) was the effect size used for the meta-analysis on depression symptoms. All effect sizes were computed using the Campbell Collaboration effect size calculator (Wilson, 2015). Data were entered in the calculator by two individuals to ensure accuracy.

Odds ratio. The *OR* is an effect size that evaluates whether the odds of a certain outcome (e.g., depression) is the same for two groups (e.g., men and women). For the meta-analysis on gender differences in major depression, the *OR* measures the ratio of the odds of major depression among women (number of depressed women divided by number of nondepressed women) to the odds of major depression among men (number of depressed men divided by number of nondepressed men). Thus, values greater than 1 indicate that women have greater odds of depression compared with men. The *OR* is different from a simple ratio of depressed women to depressed men.[6] Most studies in psychiatric epidemiology report *OR*s.

The *OR* for each independent sample was computed using either a 2×2 frequency table or the proportion depressed and total sample size for each gender. Separate effect sizes were computed for separate groups within each study (e.g., different age groups, different U.S. ethnic groups).

In synthesizing *OR*'s meta-analytically, raw *OR* effect sizes were transformed using a natural log transformation. The logged *OR*s for individual samples were weighted by the inverse of the variance, and averaged across all studies (Lipsey & Wilson, 2001). See Table 1 for a list of all raw effect sizes (not log transformed) along with corresponding study information.

Cohen's d. Cohen's *d* (Cohen, 1988) is the effect size for the standardized mean difference between two groups on a continuous variable (e.g., the mean difference between women and men on a continuous measure of levels of depression symptoms). The *d* for each sample in the depression symptom meta-analysis was computed such that positive values indicated that women reported more symptoms than men (the mean score for women minus the mean score for men, divided by the within-groups *SD*). Effect sizes of $d = 0.20$, $d = .50$, and $d = .80$ are considered to be small, medium, and large, respectively (Cohen, 1988). Separate effect sizes were computed for separate groups within each sample (e.g., different age groups, different ethnic groups). Raw effect sizes were corrected for bias (i.e., the upward-bias of effect sizes among small samples; Hedges, 1981); however, most correction factors were close to 1.0 given the large sample sizes. Although we corrected from Cohen's *d* to Hedges' *g*, we continue to refer to the results as *d* values. Effect size variances were calculated using these unbiased effect sizes. Then unbiased effect sizes for individual samples were weighted by the inverse of the variance and averaged across all studies (Lipsey & Wilson, 2001). See Table 2 for a list of all raw effect sizes (not corrected for bias) along with corresponding study information.

Data-Analytic Plan

Results were analyzed using SPSS/PASW Version 21 with macros provided by Wilson (2006). First, mean weighted effect sizes were computed for *OR* and *d*. For ease of interpretation, antilog values are reported for mean *OR*. We evaluated the homogeneity statistic (Q) to determine whether the distribution of effect sizes was heterogeneous, and thus, required further analysis. If the Q statistics associated with *OR* and *d* were significant, results were

further analyzed using a mixed-effects model to account for variability between studies (Lipsey & Wilson, 2001). The mixed-effects model assumes that variability among effects sizes beyond subject-level sampling error has both systematic components (accounted for by moderator variables) and random components (i.e., error variance). When the Q statistic is significant, this mixed-effects approach is preferable to random-effects and fixed-effects models, each of which involve untenable assumptions; the random-effects model assumes that all variability among effect sizes (beyond subject-level sampling error) is because of error and therefore not systematic, and the fixed-effects model assumes that all variability in effects sizes is accounted for by moderators. Random-effects analyses also have lower statistical power than mixed-effects models. In mixed-effects models, a random-effects variance component is estimated after accounting for moderator variables. Then the inverse variance weights are recalculated with the random variance component, and the model is refit.

Moderator variables were tested in separate analyses (one moderator as the independent variable) using an analog to analysis of variance for categorical moderator variables and an analog to regression for continuous moderator variables (i.e., weighted ordinary least squares). We used mixed-effects models with estimation via full information maximum likelihood for all moderator analyses (Wilson, 2006). Antilog values are reported for moderator analyses using *OR*.

To test developmental effects sensitively, lifetime depression diagnoses and samples with large age ranges (e.g., 18–64) were excluded for analyses with age as a moderator. Furthermore, given the focus on developmental trends, we analyzed age as both a categorical and a continuous variable. The age categories were determined based on theory, existing research, and available data. We created the categories of 13–15 and 16–19 to correspond to findings in the much-cited Hankin et al. article (Hankin et al., 1998), so that our results could be compared directly to those findings. After that, we used decades (e.g., 20–29). Below that, we formed a childhood age category that corresponded to the available data, that is, the ages at which we had data. For age as a continuous variable, we included both linear (mean-centered) and quadratic variables for age in a multiple regression, given the potential for nonlinear findings.

All analyses were run with the full data set and then rerun excluding outliers. Following procedures recommended by Tabachnick and Fidell (1996), we identified outliers as effect sizes more than 3.29 *SD*s from the mean logged *OR* and mean *d*. The value of 3.29 *SD* corresponds to $p < .001$. We report the results using the full data set, and note when results differed if outliers were excluded. Comprehensive Meta-Analysis software, Version 3 (Borenstein, Hedges, Higgins, & Rothstein, 2014) and an online program (https://vevealab.shinyapps.io/WeightFunctionModel/) were used for bias and sensitivity analyses.

---

[6] For example, consider a sample with 1,000 women and 1,000 men, where 100 women and 50 men are depressed. The ratio of depressed women to men is 2:1 (100/50). The *OR* is 2.11 ((100/900)/(50/950)). If we maintain the 2:1 ratio but increase the prevalence of depression (200 depressed women, 100 depressed men), then the *OR* increases to 2.25.

Results

Analysis for Possible Bias and File Drawer Effects

We guarded against sample bias, publication bias, and file drawer effects (Rosenthal, 1979) in several ways. First, all data sets were nationally representative, indicating that any bias in sampling is minimized.

Focus of article. Second, we coded the focus of the article to determine if the identified studies were predominantly focused on gender given the computerized database search for depression *and* gender. Notably, the majority of effect sizes were from articles that focused on depression (70% for major depression; 56% for depression symptoms), not gender (see Table 3). This indicates that most studies were published on the basis of the work on depression and not on the basis of a gender difference, so publication bias in the direction of gender differences should not be a problem.

We further tested whether effect sizes differed between articles that focused on depression rather than on gender. For the major depression meta-analysis, article focus accounted for significant variation in effect sizes (see Table 4). However, in follow-up analyses, effect sizes from articles that focused on gender were not significantly different than effect sizes from articles that focused on depression ($Q_B = 2.22$, $p = .14$) or something other than gender or depression ($Q_B = 0.91$, $p = .34$). For the depression symptom meta-analysis, article focus did not account for effect size variability. Moreover, when excluding outliers, effect sizes for articles that focused on depression and articles that focused on gender were both $d = 0.26$.

Thus, the similarity of the effect sizes for articles focused on depression and gender, combined with the small proportion of gender-focused studies, suggests that publication bias for articles finding gender differences is not a serious concern in these meta-analyses.

Unpublished data. Third, we followed up with authors to retrieve data on gender differences in depression and moderating variables when these were not reported sufficiently in the article. For many of the studies, the gender analyses were not reported in the article or were reported in little detail. For the meta-analysis on depression diagnoses, we received data from authors for 24 (37%) of the 65 articles, such that 167 (56%) of the 300 effect sizes were based on obtained rather than published data. For the meta-analysis on depression symptoms, we received data from authors for 71 (75%) of the 95 articles, such that 357 (86%) of the 413 effect sizes were based on obtained data. This protects the data from file drawer effects.

For the major depression meta-analysis, effect sizes from unpublished data ($OR = 1.83$) were significantly smaller than effect sizes from published data ($OR = 2.09$). However, the majority of effect sizes included in this meta-analysis were from unpublished data, making potential publication bias less of a concern. For the depression symptom meta-analysis, effect sizes did not differ as a function of publication status.

Funnel plot and test for asymmetry. Fourth, we used funnel plots as a visual tool to detect small-study effects. See supplemental material Figures 1 and 2 for a plot of effect size against precision (the inverse of SE) for both meta-analyses. It is important to note that the notion of "small-study effects" is in the context of relatively large nationally representative samples. The average

sample size was 5,720 (minimum = 261) for the major depression meta-analysis and 4,654 (minimum = 101) for the symptom meta-analysis. Nonetheless, we used the Begg and Mazumdar (1994) rank correlation test to evaluate asymmetry in the funnel plots. We selected this test given the skewness of the sample size variable and adequate power with the large number of effect sizes in each meta-analysis. For the major depression meta-analysis, $\tau = 0.07$, $p = .07$. For the depression symptom meta-analysis, $\tau = -0.03$, $p = .49$. Thus, neither of the tests for skewness was statistically significant, indicating no evidence of bias in the set of effect sizes, for both meta-analyses.

Sensitivity analysis. Finally, we used the Vevea and Woods Weight-Function Model for Publication Bias (Vevea & Woods, 2005). A recent review on adjusting for publication bias in meta-analysis encouraged the use of sensitivity measures (McShane, Böckenholt, & Hansen, 2016). The likelihood ratio tests (LRT) comparing the unadjusted to adjusted models (using *p* value cut points of 0.05, 0.01, and 0.001) for the major depression and depression symptom meta-analyses, respectively, were not statistically significant, $p = .07$ and $p = .32$. Although the LRT showed a marginal effect for the major depression meta-analysis, the weighted average from the unadjusted model (logged $OR = 0.67$) and the adjusted model (logged $OR = 0.62$) were quite similar. The goal of sensitivity analyses is to determine whether the results are robust to various methodological choices that were made in the process of conducting the meta-analysis. The Vevea and Woods test provides no evidence for publication bias in the sets of effects sizes in these two meta-analyses.

Description of the Samples for Each Meta-Analysis

See Table 3 for a list of descriptive information about moderator variables and other variables that describe the sample of studies for both meta-analyses. The samples of studies have similarities across the two meta-analyses. They both include mostly high-income countries and cover the life span. However, they differ in terms of the distribution of nations and year of data collection. The major depression analysis includes the most effect sizes from the U.S./Canada (46%), and the depression symptom meta-analysis includes the most effect sizes from Europe (53%). The depression symptom meta-analysis covers data collected from 1978–2014 whereas the major depression meta-analysis only includes data collected from 1991–2014.

The vast majority of effect sizes for the major depression meta-analysis were 12-month major depressive episodes based on the *DSM–IV* or *DSM–IV–TR* using a version of the CIDI. For the symptom meta-analysis, most effect sizes were based on the CES-D measure.

Magnitude of the Gender Difference in Depression

Major depression. The random-effects estimate of the weighted mean effect size for the gender difference in major depression was $OR = 1.95$, 95% confidence interval (CI) [1.88, 2.03]. The diagnosis effect size of $OR = 1.95$ is equivalent to $d = 0.37$. The random effects variance component was 0.07. The set of effect sizes using the fixed effects model was significantly heterogeneous, $Q_t(299) = 1961.63$, $p < .001$. Thus, moderator analyses were appropriate. We identified seven outlier effect sizes (2% of

Table 3

Descriptive Statistics for Moderator Variables Included in Major Depression Meta-Analysis and Depression Symptom Meta-Analysis

Moderators	Major depression (k = 300) Frequency (%) or M (SD)	Depression symptoms (k = 413) Frequency (%) or M (SD)	
Age[a]	44.81 (18.40) Range: 12–85	42.69 (22.97) Range: 8–92	
Nationality			
U.S./Canadian	137 (46%)	88 (21%)	
European	58 (19%)	219 (53%)	
Asian	28 (12%)	61 (15%)	
African	35 (8%)	0 (0%)	
Australian/New Zealander	18 (7%)	9 (2%)	
Central/South American	12 (4%)	24 (6%)	
Russian	8 (3%)	8 (2%)	
Middle Eastern	5 (2%)	4 (1%)	
GDP			
High-income	223 (74%)	330 (80%)	
Low- to middle-income	77 (26%)	83 (20%)	
Ethnicity (U.S.)			
Mixed	79 (89%)	66 (76%)	
African Americans	4 (4%)	7 (8%)	
European Americans	2 (2%)	6 (7%)	
Hispanic Americans	2 (2%)	8 (9%)	
Native Americans	2 (2%)	0 (0%)	
Asian Americans	1 (1%)	0 (0%)	
Year of data collection	2004.53 (5.52) Range: 1991–2014	2004.70 (7.25) Range: 1978–2014	
Data source			
Unpublished data	167 (56%)	337 (86%)	
Published data	133 (44%)	56 (14%)	
Focus of article			
Depression	210 (70%)	233 (56%)	
Other	62 (21%)	61 (15%)	
Gender	28 (9%)	119 (29%)	
Type of assessment			
Diagnostic interview		Symptom measure	
WMH-CIDI	135 (45%)	CES-D	198 (48%)
CIDI (not WMH- or -SF)	47 (16%)	BDI	49 (12%)
2002 World Health Survey	46 (15%)	SCL-D	36 (9%)
CIDI-SF	31 (10%)	PHQ-9	28 (7%)
AUDADIS	12 (4%)	DASS-D	11 (3%)
MINI	10 (4%)	EURO-D	11 (3%)
DISC-IV	5 (3%)	HADS-D	11 (3%)
CIS-R	4 (1%)	MHI-D	11 (3%)
DIS	4 (1%)	CDI	10 (2%)
NSA interview	3 (1%)	GDS	10 (2%)
DAWBA	2 (1%)	Other	8 (2%)
SADS	1 (<1%)	RADS	8 (2%)
		DEPS	6 (2%)
		DTS	5 (1%)
		DMI	4 (1%)
		MDI	2 (1%)
		ADMSS	1 (<1%)
		ADRS	1 (<1%)
		GHQ-D	1 (<1%)
		HDL-D	1 (<1%)
		HSCL-D	1 (<1%)
Manual			
DSM-IV/ DSM-IV-TR	265 (88%)		
DSM-III/ DSM-III-R	18 (6%)		
ICD-10	17 (6%)		
Type			
Depressive episode	215 (72%)		
Major depressive disorder	85 (28%)		
Time span			
12 months	263 (88%)		
Lifetime	23 (8%)		

Table 3 (*continued*)

Moderators	Major depression (*k* = 300) Frequency (%) or *M* (*SD*)	Depression symptoms (*k* = 413) Frequency (%) or *M* (*SD*)
6 months	9 (3%)	
1 month	4 (1%)	
Current	1 (<1%)	

Note. MDE = major depressive episode (also includes *ICD-10* depressive episode); MDD = major depressive disorder. Descriptive statistics for major depression meta-analysis reflect all 300 effect sizes, except for U.S. ethnicity (90 effect sizes). Descriptive statistics for depression symptom meta-analysis reflect all 413 effect sizes, except for U.S. ethnicity (87 effect sizes). See Tables 1 and 2 notes for an explanation of Diagnostic Interview acronyms (e.g., CIDI) and symptom Measure acronyms (e.g., CES-D).
[a] Age represents mean or midpoint for all effect sizes.

all effect sizes) that were more than 3.29 *SD*s from the mean logged *OR* (0.67 ± 0.96). After excluding these outliers, the random-effects estimate of the overall weighted mean effect size changed only slightly, *OR* = 1.94, 95% CI [1.87, 2.01].

For the benefit of U.S. policymakers, we repeated all analyses using just U.S. samples. These analyses can be found in the supplemental tables.

Depression symptoms. The random-effects estimate of the weighted mean effect size for the gender difference in depression symptoms was *d* = 0.27, 95% CI [0.26, 0.29]. The symptom effect size of *d* = 0.27 is equivalent to logged *OR* = 0.49 and *OR* = 1.64. The random effects variance component was 0.02. The set of effect sizes using the fixed effects model was significantly heterogeneous, $Q_t(412) = 9542.50$, *p* < .001. Thus, moderator analyses were appropriate. We identified three outlier effect sizes (1% of all effect sizes) that were more than 3.29 *SD*s from the mean (0.27 ± 0.47). After excluding these outliers, the random-effects estimate of the overall weighted mean effect size did not change, *d* = 0.27, 95% CI [0.25, 0.28].

Developmental Trends

As both a categorical variable (see Table 4) and continuous variable (see Table 5), age predicted variability in effect size for diagnoses and symptoms. The patterns were highly similar in both meta-analyses. See Figure 2 (for diagnoses) and Figure 3 (for symptoms) for a graphical representation of age trends.

Major depression. Effect sizes ranged from *OR* = 1.71 to *OR* = 3.02, with *OR*s >2.0 during adolescence and *OR*s between 1.71 and 2.02 in adulthood. Note that the youngest age group available for these analyses was 12 years old, making it impossible to observe the emergence of the gender difference from childhood to adolescence. When outliers were excluded, the age 13–15 *OR* decreased from 3.02 to 2.92 and the age 70 + *OR* increased from 2.02 to 2.20.

Follow-up testing with pairs of consecutive age groups indicated that the *OR*s for ages 12 (2.37), 13–15 (3.02), and 16–19 (2.69) were not statistically different (Q_B = 1.19, Q_B = 1.98, *p*s > 0.15, respectively). However, significance tests for moderators in meta-analysis tend to have low statistical power (Hedges & Pigott, 2004), which would especially be the case for age 12 when only two effect sizes were available. The *OR* at ages 16–19 (2.69) was significantly larger than the *OR* for ages 20–29 (1.93), Q_B = 43.19, *p* < .001, indicating a significant decrease in the gender

difference from adolescence to the 20s. Differences between 20 and 29 and later ages were not significant.

Depression symptoms. Effect sizes ranged from *d* = 0.09 to *d* = 0.41, peaking at ages 16–19, declining in the 20s, and staying relatively stable at roughly *d* = 0.20 after that. Removal of outliers did not change the estimates of effect sizes. In Figure 3 we present weighted effect sizes for each year in adolescence to describe in more detail the development of the gender difference in depression in adolescence (with each age having at least five effect sizes).

Follow-up testing with pairs of consecutive age groups indicated that the effect sizes for ages 8–12 (0.09), 13–15 (0.35), and 16–19 (0.41) were statistically different (Q_B = 23.01 and Q_B = 8.06, *p*s < .01, respectively), such that the effect size for each consecutive age group was significantly larger than the previous age group. The effect size for ages 20–29 (0.30) was significantly smaller than the effect size for ages 16–19, Q_B = 7.26, *p* < .01. Differences between 20 and 29 and later ages were not significant.

Nation-Level Economic Indicators

See Table 4 for income category results. See Table 5 for income inequality results.

Major depression. Income category (high vs. low to middle) was a significant predictor of effect size. Larger gender differences in depression were found in wealthier countries (*OR* = 2.00) compared with low- to middle-income countries (*OR* = 1.82). Income inequality was not a significant predictor.

Depression symptoms. Income category (high vs. low to middle) was not a significant predictor of effect size. However, when outliers were removed, the effect became marginally significant (Q_B = 3.09, *p* = .08) with smaller gender differences in high-income nations (*d* = 0.26) compared with low- to middle-income nations (*d* = 0.29). Income inequality was a significant predictor of effect size, such that larger gender differences were reported in nations with low levels of income inequality. Yet, when outliers were removed, this effect become nonsignificant (*p* = .13). Neither of these results are reliable given the sensitivity when outliers were excluded.

Nation-Level Gender Equity Indicators

See Table 5 for nation-level gender equity results.

Major depression. Contraceptive prevalence and literacy ratio both predicted variability in effect size. As the percentage of

Table 4

Weighted ANOVAs With Each Moderator Predicting Gender Differences in Major Depression (Unshaded) and Gender Differences in Depression Symptoms (Shaded)

Categorical moderators	OR	d	95% CI	k	$Q_{between}$	Q_{within}
Age group (in years)				184[a]	122.54**	231.75*
12	**2.37**	.48	[1.68, 3.37]	2		.38
13–15	**3.02**	.61	[2.76, 3.30]	24		43.13*
16–19	**2.69**	.55	[2.39, 3.03]	17		14.54
20–29	**1.93**	.36	[1.76, 2.12]	26		23.43
30–39	**1.83**	.33	[1.65, 2.03]	20		15.89
40–49	**1.71**	.30	[1.56, 1.87]	29		28.04
50–59	**1.80**	.32	[1.63, 2.00]	25		17.85
60–69	**1.79**	.32	[1.56, 2.06]	18		18.21
70+	**2.02**	.39	[1.75, 2.33]	23		70.28*
Age group (in years)				324[a]	117.90**	319.00
8–12	1.18	**.09**	[.00, .17]	13		17.52
13–15	1.89	**.35**	[.32, .39]	53		51.19
16–19	2.10	**.41**	[.37, .44]	66		84.60*
20–29	1.72	**.30**	[.23, .37]	19		9.44
30–39	1.52	**.23**	[.17, .29]	21		6.16
40–49	1.46	**.21**	[.15, .26]	24		8.27
50–59	1.41	**.19**	[.14, .25]	29		24.59
60–69	1.57	**.25**	[.21, .29]	44		66.89*
70–79	1.52	**.23**	[.18, .28]	29		29.55
80+	1.46	**.21**	[.15, .26]	26		20.80
GDP				300	4.13*	316.25
High-income	**2.00**	.38	[1.91, 2.09]	223		238.33
Low- to middle-income	**1.82**	.33	[1.69, 1.97]	77		77.92
GDP				413	1.66	405.53
High-income	1.60	**.26**	[.25, .28]	330		351.71
Low- to middle-income	1.69	**.29**	[.26, .33]	83		53.82
Ethnicity (U.S.)				89	2.24	88.34
Mixed	**2.20**	.43	[2.05, 2.36]	81		85.40
African Americans	**1.74**	.31	[1.23, 2.46]	4		2.71
European Americans	**2.23**	.44	[1.45, 3.43]	2		.04
Hispanic Americans	**1.95**	.37	[1.24, 3.07]	2		.18
Native Americans	**1.86**	.34	[1.05, 3.30]	2		.02
Ethnicity (U.S.)				87	5.24	85.92
Mixed	1.52	**.23**	[.21, .26]	66		75.20
African Americans	1.27	**.13**	[.04, .22]	7		2.59
European Americans	1.39	**.18**	[.09, .27]	6		1.69
Hispanic Americans	1.49	**.22**	[.13, .31]	8		6.44
Data Source				300	111.54**	316.34
Unpublished data	**1.83**	.33	[1.74, 1.93]	167		
Published data	**2.09**	.41	[1.98, 2.20]	133		
Data Source				413	1.21	405.99
Unpublished data	1.63	**.27**	[.26, .29]	357		367.42
Published data	1.57	**.25**	[.20, .29]	56		38.56
Focus of article				300	15.31**	319.19
Depression	**1.86**	.34	[1.77, 1.94]	210		219.21
Other	**2.21**	.44	[2.05, 2.38]	62		91.28*
Gender	**2.04**	.39	[1.80, 2.32]	28		8.69
Focus of article				413	2.37	405.52
Depression	1.60	**.26**	[.24, .28]	233		167.47
Other	1.69	**.29**	[.24, .33]	61		685.28*
Gender	1.66	**.28**	[.25, .31]	119		152.77*
Diagnostic interview				300	.85	316.59
WMH-CIDI	**1.99**	.38	[1.88, 2.11]	135		181.95*
Other	**1.92**	.36	[1.82, 2.02]	165		134.65
Symptom measure				413	18.26**	406.14
CES-D	1.57	**.25**	[.23, .28]	198		158.23
BDI	1.44	**.20**	[.15, .25]	49		23.28
Other	1.75	**.31**	[.28, .33]	166		224.63
Manual for major depression				300	2.42	315.51
DSM-IV/ DSM-IV-TR	**1.97**	.37	[1.89, 2.05]	265		282.46
DSM-III/ DSM-III-R	**1.97**	.37	[1.69, 2.30]	18		4.89
ICD-10	**1.73**	.30	[1.47, 2.03]	17		28.16*
Type for major depression				300	.82	313.80

Table 4 (*continued*)

Categorical moderators	*OR*	*d*	95% CI	*k*	$Q_{between}$	Q_{within}
Depressive episode	**1.93**	.36	[1.85, 2.02]	215		251.50[*]
Major depressive disorder	**2.01**	.38	[1.87, 2.17]	85		62.30
Time span for major depression				299	2.96	313.99
12 months	**1.96**	.37	[1.88, 2.04]	263		298.14[†]
Lifetime	**1.98**	.38	[1.72, 2.28]	23		9.78
6 months	**2.01**	.38	[1.63, 2.47]	9		5.16
1 month	**1.49**	.22	[1.09, 2.04]	4		.91

Note. Equivalent values are provided for *OR* and *d* for both meta-analyses. *OR* = weighted mean effect size; *d* = weighted mean effect size; 95% CI = 95% confidence interval for *d*; *k* = number of effect sizes for moderator or category in each moderator; $Q_{between}$ = significant values indicate that there is significant variability accounted for by the moderator; Q_{within} = significant values indicate that studies are still heterogeneous after accounting for the moderator variable; MDE = major depressive episode (also includes *ICD-10* depressive episode); ANOVA = analysis of variance; *DSM = Diagnostic and Statistical Manual for Mental Disorders*.
[a] To more sensitively test developmental effects, lifetime depression was excluded for major depression diagnoses and samples with large age ranges (e.g., 18–64) were excluded for both meta-analyses.
[*] $p < .05$.   [**] $p < .001$.   [†] $p < .10$.

women using some form of contraception increased (range = 8–84%), the effect size also increased. For literacy, the effect size increased as the ratio of the female:male adult literate population increased. Notably, there was not a range restriction for the literacy variable (ratios ranged from .31 to 1.07), which can have substantial negative skew (Else-Quest & Grabe, 2012). When outliers were excluded, executive positions had a marginal effect on effect size, such that as the percentage of executive positions held by women increased (range = 2–58%), the effect size increased. Thus, for all three indicators, greater gender equity was associated with a larger gender difference in major depression.

Depression symptoms. Contraceptive prevalence, executive positions, and the literacy ratio did not predict variation in effect size. These conclusions, however, should be qualified because the variability for all three indicators was limited.

Additional Moderators

U.S. ethnicity. In both meta-analyses, U.S. ethnicity did not account for significant variation in effect size (see Table 4).

Trends over time. As shown in Table 5, for the major depression meta-analysis, year of data collection was a significant predictor of effect size, such that gender differences were larger more recently. To better understand this pattern, we created a categorical variable for year of data collection and obtained the following effect size estimates: 1991–1996 = 1.84, *k* = 22; 1997–2002 = 1.88, *k* = 134; 2003–2008 = 1.91, *k* = 59; 2009–2014 = 2.17, *k* = 85. Thus, although the range of *OR* = 1.84 to 2.17 is not great, the positive relationship is clear. For the depression symptom meta-analysis, year of data collection did not predict variation in effect size.

Type of assessment. As shown in Table 4, diagnostic interview (WMH-CIDI v. Other) did not account for significant variation in effect sizes in the major depression meta-analysis. Symptom measure significantly predicted variation in effect size. In follow-up tests, all pairwise comparisons were significantly different from each other, such that the smallest effect size was for the BDI (*d* = 0.20) and the largest effect size was for scales other than the BDI and CES-D (*d* = 0.31).

Other major depression moderators. As shown in Table 4, manual for major depression (e.g., *DSM–IV*, *ICD-10*), type of depression (episode vs. disorder) and depression time span (1 month, 6 months, 12 months, and lifetime) did not predict effect
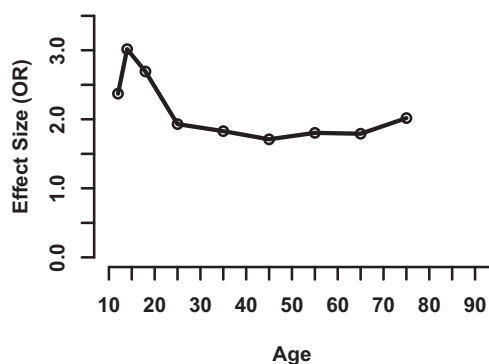
Figure 2. Effect size for gender difference in major depression across age. *Note.* Data points represent effect sizes reported in Table 4 for the following ages: 12, 13–15, 16–19, 20–29, 30–39, 40–49, 50–59, 60–69, and 70+.
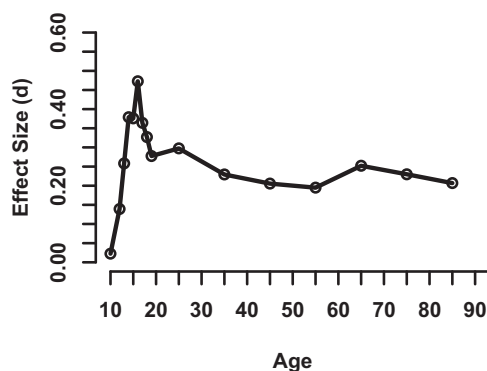
Figure 3. Effect size for gender difference in depression symptoms across age. *Note.* Data points represent effect sizes for the following ages: 8–11 (*d* = 0.02), 12 (*d* = 0.14), 13 (*d* = 0.26), 14 (*d* = 0.38), 15 (*d* = 0.38), 16 (*d* = 0.47), 17 (*d* = 0.36), 18 (*d* = 0.33), 19 (*d* = 0.28), 20–29 (*d* = 0.30), 30–39 (*d* = 0.23), 40–49 (*d* = 0.21), 50–59 (*d* = 0.19), 60–69 (*d* = 0.25), 70–79 (*d* = 0.23), and 80+ (*d* = 0.21).

Table 5

Separate Weighted OLS Regressions With Each Moderator Predicting Gender Differences in Major Depression (Unshaded) and Gender Differences in Depression Symptoms (Shaded)

Continuous moderators	β	Exp(β)	k	Q_{model}	$Q_{residual}$	R^2
Year of data collection	.14*	1.01	300	6.25*	316.48	.02
Year of data collection	.00		413	.66	405.64	.00
Age[a]			184	82.47**	221.35*	.27
Linear	−.45**	.99				
Quadratic	.35**	1.00				
Age[a]			324	40.06**	319.83	.11
Linear	−.37**					
Quadratic	.12*					
Nation-level economic indicators						
Income inequality	.09	1.01	234	2.19	252.70	.01
Nation-level economic indicators						
Income inequality	−.11*		316	3.73*	308.77	.01
Nation-level gender equity indicators						
Contraceptive prevalence	.22**	1.01	294	16.14**	308.79	.05
Executive positions	.10	1.00	256	2.65	271.35	.01
Literacy ratio	.18**	1.84	297	11.08**	313.62	.03
Nation-level gender equity indicators						
Contraceptive prevalence	.03		369	.36	362.30	.00
Executive positions	−.06		376	1.24	369.63	.00
Literacy ratio	−.03		404	.28	397.68	.00

Note. Each moderator was run in a separate regression, except for the age variables which were run in a multiple regression. Negative values indicate that there are larger reported gender differences at lower levels of the moderator and smaller reported gender differences at higher levels of the moderator. k = number of effect sizes for each moderator 95% CI = 95% confidence interval for β; Exp(β) = exponent of β; Significance for Exp(β) is not indicated because it is the same as β; Q_{model} = significant values indicate that there is significant variability accounted for by the moderator; $Q_{residual}$ = significant values indicate that studies are still heterogeneous after accounting for the moderator variable; R^2 = amount of variance accounted for by the moderator(s).
[a] = for age as a moderator, lifetime depression was excluded and samples with large age ranges (e.g., 18–64) were excluded to more sensitively test developmental effects.
† $p < .10$.   * $p < .05$.   ** $p < .001$.

size variation. However, when outliers were excluded, manual for major depression predicted significant variation in effect size, $Q_B = 7.26$, $p = .013$. Follow up analyses, excluding outliers, indicated that diagnoses using the *ICD* were significantly smaller than both diagnoses using *DSM–IV/ DSM–IV–TR* ($Q_B = 6.86$, $p < .01$) and *DSM–III/ DSM*- III-R ($Q_B = 9.60$, $p < .01$). However, the difference between $OR = 1.73$ for *ICD* and $OR = 1.97$ for both *DSM*s is not a large difference.

Discussion

The current meta-analyses advance research by synthesizing data from representative samples of more than 1.7 million women and men each, with three main goals: (a) to determine the magnitude of gender differences in diagnoses of major depression and in levels of depression symptoms; (b) to elucidate developmental trends in the magnitude of the gender difference, with the goal of identifying the age at which the gender difference in depression emerges in adolescence and whether the gender difference remains the same across adulthood; and (c) to identify other moderators of these gender differences, focusing especially on nation-level indicators of gender equity and national wealth. In the sections that follow, we highlight and discuss the findings related to each goal.

Magnitude of the Gender Difference in Depression

Overall, the *OR* was 1.95 for gender differences in diagnoses of major depression; this is the first time that this *OR* has been estimated meta-analytically and across such a large sample. For

gender differences in depression symptoms, we found $d = 0.27$; this is the first meta-analytic estimate of gender differences in symptoms based on samples across the life span.

Analyses of moderating variables revealed variations in the *magnitude* of gender differences in depression, not the *direction* of the gender difference. That is, among different subgroups, all *OR*s for diagnoses were >1.0, and all effect sizes for symptoms were positive. This emphasizes the consistency with which women have higher levels major depression and depression symptoms than men.

How do we interpret the magnitude of the gender difference? An *OR* of 1.95 is a medium, not a large, effect size, yet it is still a health disparity. Oversimplified thinking about the *OR* for gender differences in major depression diagnoses can lead to beliefs that many women are depressed and few men are. This is simply not an accurate inference with an *OR* of 1.95. For example, in a nation where 10% of women have major depression, this means that, 5.4% of men also have major depression.

One possible negative consequence of emphasizing the preponderance of women with depression is that depression becomes a female-stereotyped disorder. Such a stereotype can be harmful to both women and men. The stereotype might lead to overdiagnosis of depression in women, and, potentially, overmedication. For men, the stereotype may mean that their depression is overlooked. It is important that clinicians do not overlook depression among men, particularly because gender biases in diagnosis have been documented (Hartung & Widiger, 1998). Men may be less likely to develop depression than

women; however, this does not mean that depressed men are not distressed and impaired.

Comparison of the Diagnosis and Symptom Findings

Expressed in the Cohen's *d* metric, the two effect sizes are similar: $d = 0.37$ for major depression and $d = 0.27$ for depression symptoms. Ideally, the same samples would be included in both meta-analyses to perfectly compare these effect sizes; however, across a wide variety of nations, measures, and ages, the magnitude of the gender difference for depression symptoms and diagnoses was comparable.

We would not expect findings across the two meta-analyses to be identical given key differences between measures of depression symptoms and diagnoses of major depression. For example, the typical assessment of symptoms often represents a short period of time, such as a week, whereas diagnoses involve aggregation over longer periods, often a year (Haeffel et al., 2003). Thus, most individuals who are currently experiencing a major depressive episode will, indeed, score high on a measure of depression symptoms. However, an individual who scores in the moderate-to-low range on a measure of current depression symptoms may have experienced a major depressive episode earlier that year. Despite this difference in amount of time captured by each assessment, the magnitude of the effect for both diagnoses and symptoms was similar.

In the moderator analyses, developmental trends were also highly consistent across both meta-analyses (see Table 6 for a summary of comparisons between the symptom and diagnostic findings). However, some findings did not replicate across meta-analyses (e.g., nation-level indicators, trends over time), which may be influenced by the different set of nations and studies included in the two meta-analyses. Each of these moderator findings is discussed in the sections that follow.

The Developmental Pattern of Gender Differences in Depression

Age was the strongest predictor of effect size, compared with all other moderator variables. For both meta-analyses, the effect size

peaked in adolescence but then declined and remained stable in adulthood, a finding that has not been identified previously. The consistency of the findings across the two meta-analyses indicates that the findings are robust.

Adolescence. One of the goals of these meta-analyses was to ascertain the time course of the emerging gender difference in depression. In the major depression meta-analysis, we could not examine the emergence of the gender difference given that the youngest age in the studies was 12, when the *OR* was already 2.37. These results differ from those of Hankin and colleagues (1998), who found that that the gender gap in major depression emerged between ages 13 and 15 and then widened between ages 15 and 18. The *OR* for the 13–15 age group in our meta-analysis was already 3.02 and declined, not widened, to $OR = 2.69$ for ages 16–19.

In the symptom meta-analysis, the gender difference emerged in adolescence with a trivial gender difference for ages 8–11 (see Figure 3) and then a steep increase, reaching a peak in the gender difference at age 16. The gender difference in depression symptoms emerged somewhat earlier in adolescence in our meta-analysis ($d = .02$ for ages 8–11, $d = .14$ for age 12, $d = .26$ for age 13, $d = .38$ for age 14) compared with the Twenge and Nolen-Hoeksema (2002) meta-analysis of CDI data ($d = -.06$ for age 12, $d = .08$ for age 13, $d = .22$ for age 14). The gender difference in adolescence in our meta-analysis was also larger (largest adolescent $d = .47$ for age 16) compared with the Twenge and Nolen-Hoeksema meta-analysis (largest adolescent $d = .22$ for ages 14 and 15). The difference in findings may be because of the greater recency of many of our studies, the greater number of nations, or the inclusion of multiple measures of depression symptoms.

Taken together, our results provide powerful evidence that the gender difference in depression emerges earlier than previously thought (by *at least* age 12 for diagnoses, *at* age 12 for symptoms), which has important implications for the timing of preventive interventions.

Adulthood. In addition to clarifying the time course of the emerging gender difference in depression in adolescence, these meta-analyses also shed light on patterns of gender differences in

Table 6

Comparison of Key Findings of Diagnosis Meta-Analysis and Symptom Meta-Analysis

	Major depression	Depression symptoms
Overall effect size for gender differences	$OR = 1.95$	$d = .27$
Age trends	Significant quadratic trend. *OR* peaked at ages 13–15, declined into the 20s, and stayed stable after that.	Significant quadratic trend. No gender difference at ages 8–11. *d* peaked at age 16, declined into the 30s, and stayed stable after that.
Nation-level economic indicators		
High-income v. low- to middle-income	Larger *OR* in wealthier nations.	n.s. (significant with outliers excluded, smaller *OR* in wealthier nations).
Income inequality	n.s.	Smaller *d* in nations with greater income inequality (n.s. with outliers excluded).
Nation-level gender-equity indicators		
Contraceptive prevalence	Larger *OR* with greater contraception.	n.s.
Executive positions	n.s. (significant with outliers excluded, larger *OR* with more executive positions).	n.s.
Literacy ratio	Larger *OR* with greater female:male literacy.	n.s.
Ethnicity, U.S.	n.s.	n.s.

Note. *OR* = odds ratio; n.s. = not significant.

adulthood, an area that has been largely neglected. In both meta-analyses, the gender difference declined in early adulthood and then remained relatively stable, hovering between $OR = 1.71–2.02$ and $d = 0.19–0.30$. This pattern is a new finding and should be robust because it is based on large-scale meta-analyses and was consistent across both diagnosis and symptom measures. This finding has major implications for theories of gender differences in depression, as discussed in Theoretical Implications below.

Future empirical directions. Future research should explore how absolute levels of depression diagnoses and symptoms among men and women contribute to this pattern of a peak gender difference in adolescence, followed by a subsequent decrease and leveling off. Do men have lower depression symptoms and diagnoses in adolescence that then increase in their 20s, contributing to the observed decrease in the gender difference from adolescence to adulthood? Or do women's depression symptoms and diagnoses decrease in their 20s? Alternatively, it may be that a combination of both patterns occurs. Understanding these patterns will be important for theories of the etiology of depression and for informing prevention work. One latent growth curve analysis indicated that girls' depression symptoms accelerated early in adolescence and then leveled off, whereas boys' symptom levels accelerated in late adolescence (Salk, Petersen, Abramson, & Hyde, 2016), consistent with the first possibility above.

Theoretical implications. As noted earlier, theories guided by developmental psychopathology have focused on explaining the emergence of the gender difference in adolescence (summarized by Hyde et al., 2008b), but did not attend to development across adulthood. The strongest theory will take development into account, not only adolescent development, but also adult development. Future theoretical work will need to account not only for the peak in the magnitude of the gender difference in adolescence, but also for (a) the decline into early adulthood and (b) stability across adulthood. Here we provide examples of exciting directions in which such theorizing might go, for three factors hypothesized to be important in the development of depression: temperament, cognitive vulnerability-stress interactions, and puberty.

According to one theoretical account, *temperament*, present from infancy and early childhood, predicts later depression (summarized by Hyde et al., 2008b). In particular, individuals who are high in negative affectivity and low on positive affectivity are vulnerable to later depression. Given no gender difference in negative affectivity in infancy and childhood (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006), for temperament to account for the emergence of the gender difference in depression in adolescence requires an interaction between vulnerable temperament and some other factor, such as stress, with stress increasing dramatically in adolescence and increasing more for women than men. How, then, would such a theory account for the decline in the gender difference in the 20s and beyond? It might posit a narrowing of the gender gap in stress beginning in early adulthood. Empirical studies of developmental trends in gender differences in stress in adulthood are lacking and would be a fruitful avenue for future research.

Another theoretical account rests on *cognitive vulnerability-stress models* of depression, which have been well supported in samples of college students and adults (summarized by Hyde et al., 2008b). Research suggests that negative cognitive style may not emerge as a stable trait until ages 9.5 to 12.5, and the cognitive vulnerability-stress interaction does not become a reliable predictor of depression until ages 13.5 to 14.5, that is, in early adolescence (Cole et al., 2008). According to this model, the gender difference in depression in adolescence may be accounted for by (a) higher levels of negative cognitive style in girls than boys beginning in early adolescence; (b) higher levels of stress for girls than boys beginning in early adolescence; or (c) both. How would this theoretical framework account for the narrowing of the gender gap in depression in adulthood? One possibility is that the gender gap in negative cognitive style narrows in the 20s. The other is that the gender gap in stress narrows in early adulthood. Again, strong empirical studies of these possibilities are lacking.

Another set of theories emphasizes *biological factors* in explaining the gender difference in depression (summarized by Hyde et al., 2008b). Here we focus on *puberty* and the role of pubertal timing, which have been invoked especially to explain why the gender difference in depression appears in early adolescence. Notably, our meta-analytic findings confirm that the gender difference in depression symptoms emerges around puberty, supporting continued theorizing about the role of puberty. According to one theoretical account, early puberty is disadvantageous for girls but not boys, for outcomes such as depression (Ge, Conger, & Elder, 2001). Thus, the gender difference in depression is created at least in part by girls who go through puberty early, because of any of several processes, such as early puberty girls encountering more peer sexual harassment than boys and on-time girls (Lindberg, Grabe, & Hyde, 2007). The narrowing of the gender gap in depression in adulthood, in the early puberty account, might result from a diminution of the effects of early puberty over time (Copeland et al., 2010), especially 10 or more years later. Again, empirical data on this point are lacking, but the developmental patterns identified by our meta-analysis suggest new directions for both theory and research.

Theories in developmental psychopathology as well as sociology will also be advanced by considering why the gender difference remains relatively stable in adulthood. The following are some possible directions. First, today, at least in the United States and many other Western nations, adult women's and men's work and family roles are much more equalized than before. For example, in the 21st century women constitute 47% of the U.S. labor force, compared with 30% in 1950 (Costello, Wight, & Stone, 2003). Thus, employment is much more of a constant factor in most adult women's lives, just as it has been in men's. This may serve to level out stressors and buffers to stress across adulthood. Second, major life transitions that formerly occurred at standard ages and could be major sources of stress, no longer occur at such regular ages. Life course sociologists have called this a "de-standardization of the life course" (Bruckner & Mayer, 2005) or "disorder in the life course" (Rindfuss, Swicegood, & Rosenfeld, 1987). Today, the ages of major events such as marriage, childbirth, and divorce do not occur at the same time for all or most individuals. The result is that stressors attached to these transitions are spread out more evenly across adulthood, leading to more even rates of depression across age for both women and men, and a stable gender gap. A third possibility results from the observation that depression is a recurrent disorder (e.g., Lewinsohn, Rohde, Seeley, Klein, & Gotlib, 2000). As such, it may be that most first cases of depression occur in adolescence, and then recur at variable times in adulthood, as a result of the uneven timing of major

stressors. Again, prevalence would remain stable across adulthood for both women and men, leading to a stable gender gap, created by the original gender gap in adolescence.

Cross-National Variations

For major depression diagnoses, variability in the gender difference in depression was linked to measures of national wealth and gender equity.

Several caveats are in order before interpreting the cross-national findings. First, a different set of nations was represented in the symptom meta-analysis compared with the diagnosis meta-analysis. Second, gender equity measures were missing for some nations; thus, the analyses did not capture patterns among all nations included. Third, depression may manifest differently in different cultures (Tsai & Chentsova-Dutton, 2002). The depression measures used in the current meta-analyses used standard diagnostic interviews and symptom questionnaires that are based on Western conceptualizations of depression, as codified in the *DSM* and the *ICD*. In some cases, phrases were changed, such that the terms used to describe specific symptoms were customized to the local settings (Harkness et al., 2008). However, cultural adaptations that changed diagnostic criteria were not made and, thus, measures may fail to capture some aspects of depression that appear in other cultures (e.g., Beals et al., 2005).

Economic indicators. Following from sociological theories, we examined the relationship between gender differences in depression and a nation's wealth and income inequality. There was a marginal difference between high-income nations ($d = 0.26$) and other nations (low- to middle-income; $d = 0.29$) in the magnitude of the gender difference for symptom measures. For diagnoses, the *OR* was significantly higher in high-income nations (2.00) than in other nations (1.82). However, the difference between these *OR*s is small, and the 95% CIs overlap, suggesting that this difference is not a major effect. This absence of any substantial effect for nation-level wealth is consistent with other research (Bromet et al., 2011) indicating that economic development is not a major factor in cross-national variations in the magnitude of gender differences in depression.

Because of the feminization of poverty, we expected to observe larger gender differences in depression in nations with more income inequality. However, there was no relation to the gender effect size for diagnoses; it was negatively related for symptom measures but lost significance when outliers were excluded. Therefore, we cannot reach confident conclusions about the relationship between income inequality and the magnitude of the gender difference in depression. The findings reported here indicate that there is probably little or no relationship.

Gender equity indicators. Guided by sociological and social-structural theory, we examined the relationship between gender equity and the gender difference in depression. In the major depression meta-analysis, gender differences in depression diagnoses were larger in nations where women had more control over their reproduction, held more executive positions, and were more similar to men in literacy rates. That is, perhaps counterintuitively, gender differences were larger in nations with more gender equity, a finding that was consistent across three indicators. On the face of it, greater contraceptive prevalence should indicate women's greater control over reproduction and a lower rate of unwanted

pregnancies, which are a risk factor for depression (e.g., Mercier, Garrett, Thorp, & Siega-Riz, 2013). Yet, as noted in the introduction, some primary studies have found similar patterns to ours regarding gender differences and gender equity, for outcomes such as self-construals (Guimond et al., 2007) and emotion (Fischer, Rodriguez Mosquera, van Vianen, & Manstead, 2004).

To explain why larger psychological gender differences are sometimes found in nations with greater gender equity, we must look at the pattern according to the type of outcome measure. For *objective* measures such as mathematics performance, or for ratings of others such as mate preferences, the gender gap is smaller in more equitable nations (e.g., Else-Quest et al., 2010; Zentner & Mitura, 2012). It is in the domain of *subjective* self-ratings that gender differences are larger in more equitable nations (Wood & Eagly, 2012). Such judgments about the self require participants to construct estimates about the self. Guimond and colleagues (2007) proposed that gender differences in self-construals are larger in contexts in which individuals make intergroup social comparisons, for example, when they compare themselves to an out-group such as the other gender. Gender differences are smaller when individuals' social comparisons are made intragroup, for example, comparing themselves to other members of their own gender. For example, girls may come to see themselves as high in depression if they compare themselves to boys instead of girls. It is precisely in higher gender-equity nations, where men and women interact more (e.g., girls are more literate because they have more equal access to schooling with boys), that intergroup comparisons are likelier, leading to larger gender differences. In low-gender-equity nations, intergroup interactions are often greatly restricted, and gender differences on a variable such as depression are smaller because of mostly intragroup comparisons.

These findings of larger gender differences in nations with greater gender equity did not replicate with the meta-analysis of depression symptoms, suggesting that caution is needed in interpreting the findings for diagnoses. It will be important for future research to examine the relationship between gender equity and the gender difference in depression symptoms among a set of nations that has greater variability in gender equity.

We set out to test two other gender equity indicators—interpersonal violence against women and sexism beliefs—for their relationship to gender differences in depression across nations. We had to abandon these analyses, though, because values were available for only a minority of nations. Both variables are theoretically important, and future research is needed to improve them and make them useful in analyses such as the ones reported here.

Additional Factors Influencing the Gender Difference in Depression

Ethnicity in the United States. Guided by intersectionality theory, we examined whether effect sizes for gender differences in depression varied across U.S. ethnic groups. In both the diagnosis and symptom meta-analyses, differences among ethnic groups were not significant. Notably, gender differences were smallest for African Americans in both meta-analyses. These analyses, however, were based on small numbers of studies. Therefore, power to detect ethnic-group differences was limited. Much more work is

needed on the intersection of gender and ethnicity for depression in the United States as well as in other nations.

Trends over time. We sought to determine whether the gender difference in depression has been widening or narrowing over time. The symptom meta-analysis found no significant trend over time and the diagnosis meta-analysis found a positive trend, meaning that gender differences are growing larger. However, this effect was small, accounting for only 2% of the variance. Specifically, the *OR* increased from 1.84 in 1991–1996 to 2.17 in 2009–2014. It should be noted that the diagnosis meta-analysis did not cover the 1970s and 1980s or earlier decades, because researchers were not yet conducting studies based on nationally representative samples. The gender difference in depression should be monitored for possible changes going forward.

Implications for Policy

These meta-analytic findings can inform global health policy. Given that depression is a global health priority (World Health Organization, 2016), it is imperative to understand disparities in depression and which subgroups are most in need of services. These results suggest that women are at significantly greater risk of depression diagnoses and symptoms compared with men worldwide. Universal screening in primary care settings is imperative (O'Connor, Whitlock, Beil, & Gaynes, 2009), with a strong emphasis on screening adolescents. The emphasis on adolescents is particularly important because depression is a recurring disorder, so an episode in adolescence can predispose the individual to later episodes (e.g., Lewinsohn, Rohde, Seeley, Klein, & Gotlib, 2000). Research to identify preventive interventions is even more important (Muñoz et al., 2010). What the current meta-analyses cannot tell us is whether these interventions need to be tailored by gender. However, the magnitude of these gender findings is critically important. If global health efforts only targeted women, they would be missing a substantial proportion of depressed individuals, men.

An important clarification in regard to policy implications is that the findings of the current studies yield information only on the gender gap in depression, not on prevalence levels for either gender. For example, assuming equal numbers of men and women, an *OR* = 2.0 can result from a 9.05% prevalence in women and a 4.85% prevalence in men, or from an 18% prevalence in women and a 10.1% prevalence in men. Policymakers should monitor not only gender disparities, but also prevalence rates.

Strengths and Limitations

By synthesizing nationally representative studies with data from over 1.7 million participants spanning the globe in each of the two meta-analyses, we have provided a comprehensive quantitative review of data on gender differences in major depression diagnoses and depression symptoms across the life span. The findings represent especially strong scientific evidence because they are not based on small community or convenience samples and are instead based on representative samples with strong measurement. We also made extensive efforts to obtain data for as many national data sets as possible by conducting additional computerized searches and contacting authors. Overall, 76% of the effect sizes for the symptom meta-analysis and 56% of the effect sizes for the diagnosis meta-analysis were based

on data supplied by researchers, and not published in articles, reducing concerns about publication bias.

Despite these strengths, several limitations should be acknowledged. First, we did not have sufficient data to examine gender differences in major depression for children younger than age 12 (and we had only two samples for age 12), even though the minimum age criterion was seven. Research is needed on gender differences in major depression for children ages 7 through 12, based on nationally representative samples. The current study also highlights the need for more nationally representative data on gender differences in major depression in developing countries. The results reported here are weighted toward European and North American samples because so much more research has been conducted in those regions.

Second, focusing on large, nationally representative data sets meant that the diagnostic interviews for major depression were conducted not by skilled clinicians, but instead by trained lay interviewers, as is standard practice in these large-scale, epidemiological studies. However, evidence indicates high reliability between clinicians and lay interviewers. For example, in one methodological study, participants were given diagnostic interviews separately by a clinician and a highly trained nonclinician (Wittchen et al., 1991). Results indicated high agreement between the two; for major depressive disorder, $\kappa = 0.97$, with 99.7% agreement between the two sources (see also Brugha, Nienhuis, Bagchi, Smith, & Meltzer, 1999).

Third, both diagnostic measures and symptom questionnaire measures rely on self-reports from participants. If there are gender differences in willingness to disclose symptoms, then the resulting data may be biased. One early review concluded that the gender difference in depression is a real difference and not a measurement artifact (Weissman & Klerman, 1977). However, this issue deserves continued attention.

Fourth, our database search included only articles in English. We believe that this did not lead to the omission of nations in which English is not the predominant language for two reasons. The two meta-analyses included data from more than 90 nations. Thus, we achieved the goal of including data from a wide array of nations from all regions of the world. Moreover, mounting a study based on a nationally representative sample is a major, costly undertaking that should almost certainly result in multiple publications, at least one of them in English, so we should have detected such studies.

Conclusions

In two separate meta-analyses including nationally representative samples with over 1.7 million people each, we found evidence for a 1.95 *OR* for gender differences in major depression and a Cohen's *d* of 0.27 for gender differences in depression symptoms. Our results provide powerful evidence that the gender difference in major depression diagnoses and depression symptoms peaks in adolescence, with the gender gap in diagnoses emerging earlier than previously thought (*OR* = 2.37 at age 12). The gender gap then narrows and remains stable in adulthood, a finding that has not been identified previously and has important implications for both theory and preventive interventions. Larger gender differences in major depression were found in nations with greater gender equity and in more recent studies. The gender difference in depression represents a major health disparity, especially in adolescence, yet the magnitude of the difference indicates that depression in men should not be overlooked.

References

References marked with one asterisk (*) indicate studies included in the depression symptom meta-analysis. References marked with two asterisks (**) indicate studies included in the depression diagnosis meta-analysis.

*Aalto, A. M., Elovainio, M., Kivimäki, M., Uutela, A., & Pirkola, S. (2012). The Beck Depression Inventory and General Health Questionnaire as measures of depression in the general population: A validation study using the Composite International Diagnostic Interview as the gold standard. *Psychiatry Research, 197,* 163–171. http://dx.doi.org/10.1016/j.psychres.2011.09.008

*Abebe, D. S., Frøyland, L. R., Bakken, A., & von Soest, T. (2016). Municipal-level differences in depressive symptoms among adolescents in Norway: Results from the cross-national Ungdata study. *Scandinavian Journal of Public Health, 44,* 47–54. http://dx.doi.org/10.1177/1403494815604764

**Alaimo, K., Olson, C. M., & Frongillo, E. A. (2002). Family food insufficiency, but not low family income, is positively associated with dysthymia and suicide symptoms in adolescents. *The Journal of Nutrition, 132,* 719–725.

*Almqvist, F., Kumpulainen, K., Ikäheimo, K., Linna, S. L., Henttonen, I., Huikko, E., . . . Moilanen, I. (1999). Behavioural and emotional symptoms in 8–9-year-old children. *European Child & Adolescent Psychiatry, 8*(Suppl. 4), S7–S16. http://dx.doi.org/10.1007/PL00010698

Alsaker, F. D., Dundas, I., & Olweus, D. (1991). *A growth curve approach to the study of parental relations and depression in adolescence.* Paper presented at the meeting of the Society for Research on Child Development, April 18-20, Seattle, WA.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

*Andersen, I., Thielen, K., Nygaard, E., & Diderichsen, F. (2009). Social inequality in the prevalence of depressive disorders. *Journal of Epidemiology and Community Health, 63,* 575–581. http://dx.doi.org/10.1136/jech.2008.082719

Andrade, L., Caraveo-Anduaga, J. J., Berglund, P., Bijl, R. V., De Graaf, R., Vollebergh, W., . . . Wittchen, H. U. (2003). The epidemiology of major depressive episodes: Results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *International Journal of Methods in Psychiatric Research, 12,* 3–21. http://dx.doi.org/10.1002/mpr.138

**Andrews, G., Henderson, S., & Hall, W. (2001). Prevalence, comorbidity, disability and service utilisation. Overview of the Australian National Mental Health Survey. *The British Journal of Psychiatry, 178,* 145–153. http://dx.doi.org/10.1192/bjp.178.2.145

Angst, J., Gamma, A., Gastpar, M., Lépine, J. P., Mendlewicz, J., Tylee, A., & the Depression Research in European Society Study. (2002). Gender differences in depression. Epidemiological findings from the European DEPRES I and II studies. *European Archives of Psychiatry and Clinical Neuroscience, 252,* 201–209. http://dx.doi.org/10.1007/s00406-002-0381-6

APA Publications and Communications Board Working Group on Journal Article Reporting Standards. (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist, 63,* 839–851. http://dx.doi.org/10.1037/0003-066X.63.9.839

Arokiasamy, P., Parasuraman, S., Sekher, T. V., & Lhungdim, H. (2013). *India: Study on global AGEing and adult health (SAGE), Wave 1.* International Institute for Population Sciences. Geneva, Switzerland: World Health Organization.

Avenevoli, S., Knight, E., Kessler, R. C., & Merikangas, K. R. (2008). Epidemiology of depression in children and adolescents. In J. R. Z. Abela & B. L. Hankin (Eds.), *Handbook of depression in children and adolescents* (pp. 6–32). New York, NY: Guilford Press.

**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.e2. http://dx.doi.org/10.1016/j.jaac.2014.10.010

Barnes, D. M., Keyes, K. M., & Bates, L. M. (2013). Racial differences in depression in the United States: How do subgroup analyses inform a paradox? *Social Psychiatry and Psychiatric Epidemiology, 48,* 1941–1949. http://dx.doi.org/10.1007/s00127-013-0718-7

**Beals, J., Manson, S. M., Whitesell, N. R., Mitchell, C. M., Novins, D. K., Simpson, S., & Spicer, P. (2005). Prevalence of major depressive episode in two American Indian reservation populations: Unexpected findings with a structured interview. *The American Journal of Psychiatry, 162,* 1713–1722. http://dx.doi.org/10.1176/appi.ajp.162.9.1713

Bebbington, P. E. (1996). The origins of sex differences in depressive disorder: Bridging the gap. *International Review of Psychiatry, 8,* 295–332. http://dx.doi.org/10.3109/09540269609051547

Bebbington, P. E., Dunn, G., Jenkins, R., Lewis, G., Brugha, T., Farrell, M., & Meltzer, H. (1998). The influence of age and sex on the prevalence of depressive conditions: Report from the National Survey of Psychiatric Morbidity. *Psychological Medicine, 28,* 9–19. http://dx.doi.org/10.1017/S0033291797006077

Bech, P., Rasmussen, N. A., Olsen, L. R., Noerholm, V., & Abildgaard, W. (2001). The sensitivity and specificity of the Major Depression Inventory, using the Present State Examination as the index of diagnostic validity. *Journal of Affective Disorders, 66,* 159–164. http://dx.doi.org/10.1016/S0165-0327(00)00309-8

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II.* San Antonio, TX: Psychological Corporation.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571. http://dx.doi.org/10.1001/archpsyc.1961.01710120031004

Beekman, A. T., Deeg, D. J., Van Limbeek, J., Braam, A. W., De Vries, M. Z., & Van Tilburg, W. (1997). Criterion validity of the Center for Epidemiologic Studies Depression Scale (CES-D): Results from a community-based sample of older subjects in The Netherlands. *Psychological Medicine, 27,* 231–235. http://dx.doi.org/10.1017/S003329179600351091

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50,* 1088–1101. http://dx.doi.org/10.2307/2533446

*Bélanger, R. E., Akre, C., Berchtold, A., & Michaud, P. A. (2011). A U-shaped association between intensity of Internet use and adolescent health. *Pediatrics, 127,* e330–e335. http://dx.doi.org/10.1542/peds.2010-1235

Belle, D. (1990). Poverty and women's mental health. *American Psychologist, 45,* 385–389. http://dx.doi.org/10.1037/0003-066X.45.3.385

Belle, D., & Doucet, J. (2003). Poverty, inequality, and discrimination as sources of depression among U.S. women. *Psychology of Women Quarterly, 27,* 1–12. http://dx.doi.org/10.1111/1471-6402.00090

**Bijl, R. V., Ravelli, A., & van Zessen, G. (1998). Prevalence of psychiatric disorder in the general population: Results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology, 33,* 587–595. http://dx.doi.org/10.1007/s001270050098

Bor, W., Dean, A. J., Najman, J., & Hayatbakhsh, R. (2014). Are child and adolescent mental health problems increasing in the 21st century? A systematic review. *Australian and New Zealand Journal of Psychiatry, 48,* 606–616. http://dx.doi.org/10.1177/0004867414533834

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. (2014). *Comprehensive meta-analysis* (3.3.070 ed.). Englewood, NJ: Biostat, Inc.

**Boyd, A., Van de Velde, S., Vilagut, G., de Graaf, R., O'Neill, S., Florescu, S., . . . the EU-WMH Investigators. (2015). Gender differences in mental disorders and suicidality in Europe: Results from a large cross-sectional population-based study. *Journal of Affective Disorders, 173,* 245–254. http://dx.doi.org/10.1016/j.jad.2014.11.002

*Bracke, P. (1998). Sex differences in the course of depression: Evidence from a longitudinal study of a representative sample of the Belgian population. *Social Psychiatry and Psychiatric Epidemiology, 33,* 420–429. http://dx.doi.org/10.1007/s001270050075

Breslau, J., Kendler, K. S., Su, M., Gaxiola-Aguilar, S., & Kessler, R. C. (2005). Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States. *Psychological Medicine, 35,* 317–327. http://dx.doi.org/10.1017/S0033291704003514

**Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., . . . Kessler, R. C. (2011). Cross-national epidemiology of *DSM–IV* major depressive episode. *BMC Medicine, 9,* 90. http://dx.doi.org/10.1186/1741-7015-9-90

**Bromet, E. J., Gluzman, S. F., Paniotto, V. I., Webb, C. P., Tintle, N. L., Zakhozha, V., . . . Schwartz, J. E. (2005). Epidemiology of psychiatric and alcohol disorders in Ukraine: Findings from the Ukraine World Mental Health survey. *Social Psychiatry and Psychiatric Epidemiology, 40,* 681–690. http://dx.doi.org/10.1007/s00127-005-0927-9

Brückner, H., & Mayer, K. U. (2005). De-standardization of the life course: What it might mean? And if it means anything, whether it actually took place? *Advances in Life Course Research, 9,* 27–53. http://dx.doi.org/10.1016/S1040-2608(04)09002-1

Brugha, T. S., Nienhuis, F., Bagchi, D., Smith, J., & Meltzer, H. (1999). The survey form of SCAN: The feasibility of using experienced lay survey interviewers to administer a semi-structured systematic clinical assessment of psychotic and non-psychotic disorders. *Psychological Medicine, 29,* 703–711. http://dx.doi.org/10.1017/S003329179900834X

*Bushman, B. J., Moeller, S. J., Konrath, S., & Crocker, J. (2012). Investigating the link between liking versus wanting self-esteem and depression in a nationally representative sample of American adults. *Journal of Personality, 80,* 1453–1469. http://dx.doi.org/10.1111/j.1467-6494.2012.00781.x

*Cardozo, B. L., Bilukha, O. O., Gotway, C. A., Wolfe, M. I., Gerber, M. L., & Anderson, M. (2005). Report from the CDC: Mental health of women in postwar Afghanistan. *Journal of Women's Health, 14,* 285–293. http://dx.doi.org/10.1089/jwh.2005.14.285

*Cater, Å. K., Miller, L. E., Howell, K. H., & Graham-Bermann, S. A. (2015). Childhood exposure to intimate partner violence and adult mental health problems: Relationships with gender and age of exposure. *Journal of Family Violence, 30,* 875–886. http://dx.doi.org/10.1007/s10896-015-9703-0

Center for Behavioral Health Statistics and Quality. (2014). *2013 National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

*Chan, A., Malhotra, C., Malhotra, R., & Ostbye, T. (2011). Living arrangements, social networks and depressive symptoms among older men and women in Singapore. *International Journal of Geriatric Psychiatry, 26,* 630–639. http://dx.doi.org/10.1002/gps.2574

**Chan Chee, C., Gourier-Fréry, C., Guignard, R., & Beck, F. (2011). The current state of mental health surveillance in France. *Santé Publique, 23*(Suppl. 6), S13–S29.

**Cho, M. J., Chang, S. M., Lee, Y. M., Bae, A., Ahn, J. H., Son, J., . . . Kim, B. S. (2010). Prevalence of *DSM–IV* major mental disorders among Korean adults: A 2006 National Epidemiologic Survey (KECA-R). *Asian Journal of Psychiatry, 3,* 26–30. http://dx.doi.org/10.1016/j.ajp.2010.01.009

Cho, M. J., Hahm, B. J., Suh, D. W., Hong, J. P., Bae, J. N., Kim, J. K., . . . Cho, S. J. (2002). Development of a Korean version of the composite international diagnostic interview (K-CIDI). *Journal of Korean Neuropsychiatric Association, 41,* 123–137.

**Cho, M. J., Kim, J. K., Jeon, H. J., Suh, T., Chung, I. W., Hong, J. P., . . . Hahm, B. J. (2007). Lifetime and 12-month prevalence of *DSM–IV* psychiatric disorders among Korean adults. *Journal of Nervous and Mental Disease, 195,* 203–210. http://dx.doi.org/10.1097/01.nmd.0000243826.40732.45

Cicchetti, D., & Rogosch, F. A. (2002). A developmental psychopathology perspective on adolescence. *Journal of Consulting and Clinical Psychology, 70,* 6–20. http://dx.doi.org/10.1037/0022-006X.70.1.6

Clark, T., Fleming, T., Bullen, P., Crengle, S., Denny, S., Dyson, B., . . . Rossen, F. (2013). *Youth'12 Prevalence Tables: The health and wellbeing of New Zealand secondary school students in 2012*. Auckland, New Zealand: University of Auckland.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cole, D. A., Ciesla, J. A., Dallaire, D. H., Jacquez, F. M., Pineda, A. Q., Lagrange, B., . . . Felton, J. W. (2008). Emergence of attributional style and its relation to depressive symptoms. *Journal of Abnormal Psychology, 117,* 16–31. http://dx.doi.org/10.1037/0021-843X.117.1.16

Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist, 64,* 170–180. http://dx.doi.org/10.1037/a0014564

*Collins, A. L., Glei, D. A., & Goldman, N. (2009). The role of life satisfaction and depressive symptoms in all-cause mortality. *Psychology and Aging, 24,* 696–702. http://dx.doi.org/10.1037/a0016777

Copeland, W., Shanahan, L., Miller, S., Costello, E. J., Angold, A., & Maughan, B. (2010). Do the negative effects of early pubertal timing on adolescent girls continue into young adulthood? *The American Journal of Psychiatry, 167,* 1218–1225. http://dx.doi.org/10.1176/appi.ajp.2010.09081190

Costello, C. B., Wight, V. R., & Stone, A. J. (2003). *The American woman 2003–2004*. New York, NY: Palgrave/Macmillan.

**Coyne, J. C., & Marcus, S. C. (2006). Health disparities in care for depression possibly obscured by the clinical significance criterion. *The American Journal of Psychiatry, 163,* 1577–1579. http://dx.doi.org/10.1176/ajp.2006.163.9.1577

*Crimmins, E. M., Kim, J. K., & Solé-Auró, A. (2011). Gender differences in health: Results from SHARE, ELSA and HRS. *European Journal of Public Health, 21,* 81–91. http://dx.doi.org/10.1093/eurpub/ckq022

Cross-National Collaborative Group. (1992). The changing rate of major depression. Cross-national comparisons. *Journal of the American Medical Association, 268,* 3098–3105. http://dx.doi.org/10.1001/jama.1992.03490210080039

Cuijpers, P., de Graaf, R., & van Dorsselaer, S. (2004). Minor depression: Risk profiles, functional disability, health care use and risk of developing major depression. *Journal of Affective Disorders, 79,* 71–79. http://dx.doi.org/10.1016/S0165-0327(02)00348-8

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine, 11,* 126. http://dx.doi.org/10.1186/1741-7015-11-126

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,* 21–27. http://dx.doi.org/10.1001/archpsyc.57.1.21

**Danielson, C. K., de Arellano, M. A., Kilpatrick, D. G., Saunders, B. E., & Resnick, H. S. (2005). Child maltreatment in depressed adolescents: Differences in symptomatology based on history of abuse. *Child Maltreatment, 10,* 37–48. http://dx.doi.org/10.1177/1077559504271630

**de Graaf, R., ten Have, M., van Gool, C., & van Dorsselaer, S. (2012). Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. *Social Psychiatry and Psychiatric Epidemiology, 47,* 203–213. http://dx .doi.org/10.1007/s00127-010-0334-8

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An Outpatient Psychiatric Rating Scale—Preliminary report. *Psychopharmacology Bulletin, 9,* 13–28.

*de Wit, L. M., van Straten, A., van Herten, M., Penninx, B. W., & Cuijpers, P. (2009). Depression and body mass index, a u-shaped association. *BMC Public Health, 9,* 14. http://dx.doi.org/10.1186/1471-2458-9-14

*Dooley, B., Fitzgerald, A., & Giollabhui, N. M. (2015). The risk and protective factors associated with depression and anxiety in a national sample of Irish adolescents. *Irish Journal of Psychological Medicine, 32,* 93–105. http://dx.doi.org/10.1017/ipm.2014.83

Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist, 54,* 408–423. http://dx.doi.org/10.1037/0003-066X.54.6.408

Else-Quest, N. M., & Grabe, S. (2012). The political is personal: Measurement and application of nation-level indicators of gender equity in psychological research. *Psychology of Women Quarterly, 36,* 131–144. http://dx.doi.org/10.1177/0361684312441592

Else-Quest, N. M., & Hyde, J. S. (2016a). Intersectionality in quantitative psychological research: I. *Theoretical and epistemological issues. Psychology of Women Quarterly, 40,* 155–170. http://dx.doi.org/10.1177/0361684316629797

Else-Quest, N. M., & Hyde, J. S. (2016b). Intersectionality in quantitative psychological research: II. *Methods and techniques. Psychology of Women Quarterly, 40,* 319–336. http://dx.doi.org/10.1177/0361684316647953

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin, 132,* 33–72. http://dx.doi.org/10.1037/0033-2909.132.1.33

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136,* 103–127. http://dx.doi.org/10.1037/a0018053

Endicott, J., & Spitzer, R. L. (1978). A diagnostic interview: The schedule for affective disorders and schizophrenia. *Archives of General Psychiatry, 35,* 837–844. http://dx.doi.org/10.1001/archpsyc .1978.01770310043002

*European Social Survey. (2015). *Measuring and Reporting on Europeans' Wellbeing: Findings from the European Social Survey.* Retrieved from www.esswellbeingmatters.org

*Everson-Rose, S. A., House, J. S., & Mero, R. P. (2004). Depressive symptoms and mortality risk in a national sample: Confounding effects of health status. *Psychosomatic Medicine, 66,* 823–830. http://dx.doi .org/10.1097/01.psy.0000145903.75432.1f

**Farbstein, I., Mansbach-Kleinfeld, I., Levinson, D., Goodman, R., Levav, I., Vograft, I., . . . Apter, A. (2010). Prevalence and correlates of mental disorders in Israeli adolescents: Results from a national mental health survey. *Journal of Child Psychology and Psychiatry, 51,* 630–639. http://dx.doi.org/10.1111/j.1469-7610.2009.02188.x

Fergusson, D. M., Horwood, L. J., Ridder, E. M., & Beautrais, A. L. (2005). Subthreshold depression in adolescence and mental health outcomes in adulthood. *Archives of General Psychiatry, 62,* 66–72. http:// dx.doi.org/10.1001/archpsyc.62.1.66

*Ferketich, A. K., Schwartzbaum, J. A., Frid, D. J., & Moeschberger, M. L. (2000). Depression as an antecedent to heart disease among women and men in the NHANES I study. National Health and Nutrition Examination Survey. *Archives of Internal Medicine, 160,* 1261–1268. http://dx .doi.org/10.1001/archinte.160.9.1261

Ferrari, A. J., Somerville, A. J., Baxter, A. J., Norman, R., Patten, S. B., Vos, T., & Whiteford, H. A. (2013). Global variation in the prevalence and incidence of major depressive disorder: A systematic review of the epidemiological literature. *Psychological Medicine, 43,* 471–481. http:// dx.doi.org/10.1017/S0033291712001511

Fischer, A. H., Rodriguez Mosquera, P. M., van Vianen, A. E. M., & Manstead, A. S. R. (2004). Gender and culture differences in emotion. *Emotion, 4,* 87–94. http://dx.doi.org/10.1037/1528-3542.4.1.87

*Fleiz Bautista, C., Villatoro Velázquez, J., Mora Icaza, M. E., Moreno López, M., Gutiérrez López, M. L., & Oliva Robles, N. (2012). Sociodemographic and personal factors related to depressive symptomatology in the Mexican population aged 12 to 65. *Revista Brasileira de Psiquiatria, 34,* 395–404. http://dx.doi.org/10.1016/j.rbp.2012.03.004

*Fleming, T. M., Clark, T., Denny, S., Bullen, P., Crengle, S., Peiris-John, R., . . . Lucassen, M. (2014). Stability and change in the mental health of New Zealand secondary school students 2007–2012: Results from the national adolescent health surveys. *Australian and New Zealand Journal of Psychiatry, 48,* 472–480. http://dx.doi.org/10.1177/00048674 13514489

**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 Academy of Child & Adolescent Psychiatry, 42,* 1203–1211. http://dx.doi.org/10.1097/00004583-200310000-00011

**Gabilondo, A., Rojas-Farreras, S., Vilagut, G., Haro, J. M., Fernández, A., Pinto-Meza, A., & Alonso, J. (2010). Epidemiology of major depressive episode in a southern European country: Results from the ESEMeD-Spain project. *Journal of Affective Disorders, 120*(1–3), 76–85. http://dx.doi.org/10.1016/j.jad.2009.04.016

*Gault-Sherman, M., Silver, E., & Sigfúsdóttir, I. D. (2009). Gender and the associated impairments of childhood sexual abuse: A national study of Icelandic youth. *Social Science & Medicine, 69,* 1515–1522. http:// dx.doi.org/10.1016/j.socscimed.2009.08.037

**Gavin, A. R., Chae, D. H., & Takeuchi, D. (2009). Association between family structure in childhood and lifetime depressive disorder in adulthood among a nationally representative sample of Blacks. *General Hospital Psychiatry, 31,* 484–493. http://dx.doi.org/10.1016/j .genhosppsych.2009.06.002

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,* 404–417. http://dx.doi.org/10.1037/0012-1649.37.3.404

*Gettler, L. T., & Oka, R. C. (2016). Are testosterone levels and depression risk linked based on partnering and parenting? Evidence from a large population-representative study of U.S. men and women. *Social Science & Medicine, 163,* 157–167.

Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well-Being Assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, 41,* 645–655. http://dx.doi.org/10.1111/j.1469-7610.2000.tb02345.x

**Goodwin, R. D., & Gotlib, I. H. (2004). Gender differences in depression: The role of personality factors. *Psychiatry Research, 126,* 135–142. http://dx.doi.org/10.1016/j.psychres.2003.12.024

*,**Graham, K., Massak, A., Demers, A., & Rehm, J. (2007). Does the association between alcohol consumption and depression depend on how they are measured? *Alcoholism, Clinical and Experimental Research, 31,* 78–88. http://dx.doi.org/10.1111/j.1530-0277.2006.00274.x

**Grant, B. F. (1995). Comorbidity between *DSM–IV* drug use disorders and major depression: Results of a national survey of adults. *Journal of Substance Abuse, 7,* 481–497. http://dx.doi.org/10.1016/0899-3289 (95)90017-9

Grant, B. F., Dawson, D. A., & Hasin, D. S. (2001). *The Alcohol Use Disorder and Associated Disabilities Interview Schedule–DSM–IV Ver-*

sion. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.

*Guarnaccia, P. J., & Worobey, J. L. (1991). The impact of marital status and employment status on depressive affect for Hispanic Americans. *Journal of Community Psychology, 19,* 136–149. http://dx.doi.org/10.1002/1520-6629(199104)19:2<136::AID-JCOP2290190205>3.0.CO;2-6

*Gudmundsdottir, G., & Vilhjalmsson, R. (2010). Group differences in outpatient help-seeking for psychological distress: Results from a national prospective study of Icelanders. *Scandinavian Journal of Public Health, 38,* 160–167. http://dx.doi.org/10.1177/1403494809357357

Guimond, S., Branscombe, N. R., Brunot, S., Buunk, A. P., Chatard, A., Désert, M., . . . Yzerbyt, V. (2007). Culture, gender, and the self: Variations and impact of social comparison processes. *Journal of Personality and Social Psychology, 92,* 1118–1134. http://dx.doi.org/10.1037/0022-3514.92.6.1118

**Gureje, O., Uwakwe, R., Oladeji, B., Makanjuola, V. O., & Esan, O. (2010). Depression in adult Nigerians: Results from the Nigerian Survey of Mental Health and Well-being. *Journal of Affective Disorders, 120,* 158–164. http://dx.doi.org/10.1016/j.jad.2009.04.030

**Haarasilta, L., Marttunen, M., Kaprio, J., & Aro, H. (2001). The 12-month prevalence and characteristics of major depressive episode in a representative nationwide sample of adolescents and young adults. *Psychological Medicine, 31,* 1169–1179. http://dx.doi.org/10.1017/S0033291701004573

Haeffel, G. J., Abramson, L. Y., Voelz, Z. R., Metalsky, G. I., Halberstadt, L., Dykman, B. M., . . . Alloy, L. B. (2003). Cognitive vulnerability to depression and lifetime history of Axis I psychopathology: A comparison of negative cognitive styles (CSQ) and dysfunctional attitudes (DAS). *Journal of Cognitive Psychotherapy, 17,* 3–22. http://dx.doi.org/10.1891/jcop.17.1.3.58269

Hammarström, A., Lehti, A., Danielsson, U., Bengs, C., & Johansson, E. E. (2009). Gender-related explanatory models of depression: A critical evaluation of medical articles. *Public Health, 123,* 689–693. http://dx.doi.org/10.1016/j.puhe.2009.09.010

Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology, 107,* 128–140. http://dx.doi.org/10.1037/0021-843X.107.1.128

*Hardie, J. H. (2014). The consequences of unrealized occupational goals in the transition to adulthood. *Social Science Research, 48,* 196–211. http://dx.doi.org/10.1016/j.ssresearch.2014.06.006

Harkness, J., Pennell, B. E., Villar, A., Gebler, N., Aguilar-Gaxiola, S., & Bilgen, I. (2008). Translation procedures and translation assessment in the World Mental Health Survey Initiative. In R. C. Kessler & T. B. Üstün (Eds.), *The WHO World Mental Health Surveys: Global perspectives on the epidemiology of mental disorders* (pp. 91–113). New York, NY: Cambridge University Press.

*Haroz, E. E., Ybarra, M. L., & Eaton, W. W. (2014). Psychometric evaluation of a Self-Report Scale to measure adolescent depression: The CESDR-10 in two national adolescent samples in the United States. *Journal of Affective Disorders, 158,* 154–160. http://dx.doi.org/10.1016/j.jad.2014.02.009

Hartung, C. M., & Widiger, T. A. (1998). Gender differences in the diagnosis of mental disorders: Conclusions and controversies of the *DSM–IV*. *Psychological Bulletin, 123,* 260–278. http://dx.doi.org/10.1037/0033-2909.123.3.260

**Hasin, D. S., Goodwin, R. D., Stinson, F. S., & Grant, B. F. (2005). Epidemiology of major depressive disorder: Results from the National Epidemiologic Survey on Alcoholism and Related Conditions. *Archives of General Psychiatry, 62,* 1097–1106. http://dx.doi.org/10.1001/archpsyc.62.10.1097

*Haukkala, A., Konttinen, H., Uutela, A., Kawachi, I., & Laatikainen, T. (2009). Gender differences in the associations between depressive symptoms, cardiovascular diseases, and all-cause mortality. *Annals of Epidemiology, 19,* 623–629. http://dx.doi.org/10.1016/j.annepidem.2009.01.010

*Häuser, W., Schmutzer, G., Brähler, E., Schiltenwolf, M., & Hilbert, A. (2014). The impact of body weight and depression on low back pain in a representative population sample. *Pain Medicine, 15,* 1316–1327. http://dx.doi.org/10.1111/pme.12458

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128. http://dx.doi.org/10.2307/1164588

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269,* 41–45. http://dx.doi.org/10.1126/science.7604277

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9,* 426–445. http://dx.doi.org/10.1037/1082-989X.9.4.426

Hopcroft, R. L., & Bradley, D. B. (2007). The sex difference in depression across 29 countries. *Social Forces, 85,* 1483–1507. http://dx.doi.org/10.1353/sof.2007.0071

*Huang, Y.-T., & Chen, Y.-W. (2015). Risk factors for depressed mood in a Taiwanese school-based sample of adolescents: Does spirituality have protective effects? *British Journal of Social Work, 45,* 2020–2037. http://dx.doi.org/10.1093/bjsw/bcu016

*Hwang, J. M., Cheong, P. H., & Feeley, T. H. (2009). Being young and feeling blue in Taiwan: Examining adolescent depressive mood and online and offline activities. *New Media & Society, 11,* 1101–1121. http://dx.doi.org/10.1177/1461444809341699

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008a). Diversity. Gender similarities characterize math performance. *Science, 321,* 494–495. http://dx.doi.org/10.1126/science.1160364

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

*Inaba, A., Thoits, P. A., Ueno, K., Gove, W. R., Evenson, R. J., & Sloan, M. (2005). Depression in the United States and Japan: Gender, marital status, and SES patterns. *Social Science & Medicine, 61,* 2280–2292. http://dx.doi.org/10.1016/j.socscimed.2005.07.014

Instituto Nacional de Salud Publica (INSP). (2014). *Mexico: Study on global AGEing and adult health (SAGE), Wave 1*. Geneva, Switzerland: World Health Organization.

*Ismayilova, L., Hmoud, O., Alkhasawneh, E., Shaw, S., & El-Bassel, N. (2013). Depressive symptoms among Jordanian youth: Results of a national survey. *Community Mental Health Journal, 49,* 133–140. http://dx.doi.org/10.1007/s10597-012-9529-7

*Jaddou, H. Y., Batieha, A. M., Khader, Y. S., Kanaan, S. H., El-Khateeb, M. S., & Ajlouni, K. M. (2012). Depression is associated with low levels of 25-hydroxyvitamin D among Jordanian adults: Results from a national population survey. *European Archives of Psychiatry and Clinical Neuroscience, 262,* 321–327. http://dx.doi.org/10.1007/s00406-011-0265-8

*Jang, S. N., Kawachi, I., Chang, J., Boo, K., Shin, H. G., Lee, H., & Cho, S. I. (2009). Marital status, gender, and depression: Analysis of the baseline survey of the Korean Longitudinal Study of Ageing (KLoSA). *Social Science & Medicine, 69,* 1608–1615. http://dx.doi.org/10.1016/j.socscimed.2009.09.007

**Joe, S., Baser, R. S., Neighbors, H. W., Caldwell, C. H., & Jackson, J. S. (2009). 12-month and lifetime prevalence of suicide attempts among black adolescents in the National Survey of American Life. *Journal of the American Academy of Child & Adolescent Psychiatry, 48,* 271–282. http://dx.doi.org/10.1097/CHI.0b013e318195bccf

*Kaji, T., Mishima, K., Kitamura, S., Enomoto, M., Nagase, Y., Li, L., . . . Uchiyama, M. (2010). Relationship between late-life depression and life stressors: Large-scale cross-sectional study of a representative sample of the Japanese general population. *Psychiatry and Clinical Neurosciences, 64,* 426–434. http://dx.doi.org/10.1111/j.1440-1819.2010.02097.x

*Karadag, H., Oner, O., Karaoglan, A., Orsel, S., Demir, A. U., Firat, H., . . . Itil, O. (2014). Body mass index and sexual dysfunction in males and females in a population study. *Klinik Psikofarmakoloji Bülteni/Bulletin of Clinical Psychopharmacology, 24,* 76–83.

Kendall, P. C., Hollon, S. D., Beck, A. T., Hammen, C. L., & Ingram, R. E. (1987). Issues and recommendations regarding use of the Beck Depression Inventory. *Cognitive Therapy and Research, 11,* 289–299. http://dx.doi.org/10.1007/BF01186280

Kessler, R. C., & Ustun, T. B. (Eds.). (2008). *The WHO World Mental Health Surveys: Global perspectives on the epidemiology of mental disorders.* New York, NY: Cambridge University Press.

Kessler, R. C., Andrews, G., Mroczek, D., Ustun, B., & Wittchen, H. U. (1998). The World Health Organization Composite International Diagnostic Interview Short Form (CIDI-SF). *International Journal of Methods in Psychiatric Research, 7,* 171–185. http://dx.doi.org/10.1002/mpr.47

**Kessler, R. C., Birnbaum, H., Bromet, E., Hwang, I., Sampson, N., & Shahly, V. (2010). Age differences in major depression: Results from the National Comorbidity Survey Replication (NCS-R). *Psychological Medicine, 40,* 225–237. http://dx.doi.org/10.1017/S0033291709990213

**Kessler, R. C., McGonagle, K. A., Swartz, M., Blazer, D. G., & Nelson, C. B. (1993). Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. *Journal of Affective Disorders, 29,* 85–96. http://dx.doi.org/10.1016/0165-0327(93)90026-G

Kessler, R. C., & Ustün, 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,* 93–121. http://dx.doi.org/10.1002/mpr.168

**Kienja, A., Piotrowski, P., Adamowski, T., Moskalewicz, J., Wciórka, J., Stokwiszewski, J., . . . Kessler, R. (2015). The prevalence of common mental disorders in the population of adult Poles by sex and age structure - an EZOP Poland study. *Psychiatria Polska, 49,* 15–27. http://dx.doi.org/10.12740/PP/30811

*Kim, H. S. (2016). Sexual debut and mental health among South Korean adolescents. *Journal of Sex Research, 53,* 313–320. http://dx.doi.org/10.1080/00224499.2015.1055855

**Kim, J. H., Cho, M. J., Hong, J. P., Bae, J. N., Cho, S. J., Hahm, B. J., . . . Chang, S. M. (2015). Gender differences in depressive symptom profile: Results from nationwide general population surveys in Korea. *Journal of Korean Medical Science, 30,* 1659–1666. http://dx.doi.org/10.3346/jkms.2015.30.11.1659

Klein, D. N., Shankman, S. A., Lewinsohn, P. M., & Seeley, J. R. (2009). Subthreshold depressive disorder in adolescents: Predictors of escalation to full-syndrome depressive disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 48,* 703–710. http://dx.doi.org/10.1097/CHI.0b013e3181a56606

*Kliem, S., Mößle, T., Zenger, M., & Brähler, E. (2014). Reliability and validity of the Beck Depression Inventory-Fast Screen for medical patients in the general German population. *Journal of Affective Disorders, 156,* 236–239. http://dx.doi.org/10.1016/j.jad.2013.11.024

Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin, 125,* 470–500. http://dx.doi.org/10.1037/0033-2909.125.4.470

*Klomek, A. B., Sourander, A., Niemelä, S., Kumpulainen, K., Piha, J., Tamminen, T., . . . Gould, M. S. (2009). Childhood bullying behaviors as a risk for suicide attempts and completed suicides: A population-based birth cohort study. *Journal of the American Academy of Child &

Adolescent Psychiatry, 48, 254–261. http://dx.doi.org/10.1097/CHI.0b013e318196b91f

*Kocalevent, R. D., Hinz, A., & Brähler, E. (2013). Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry, 35,* 551–555. http://dx.doi.org/10.1016/j.genhosppsych.2013.04.006

*Kokkevi, A., & Fotiou, A. (2009). *The psychosocial module.* In B. Hibell, U. Guttormsson, S. Ahlstrom, O. Balakireva, T. Bjarnason, A. Kokkevi, & L. Kraus (Eds.), The 2007 ESPAD report: Substance use among students in 35 European countries. Stockholm, Sweden: The Swedish Council for Information on Alcohol and other Drugs (CAN).

*Kokkevi, A., Rotsika, V., Arapaki, A., & Richardson, C. (2011). Changes in associations between psychosocial factors and suicide attempts by adolescents in Greece from 1984 to 2007. *European Journal of Public Health, 21,* 694–698. http://dx.doi.org/10.1093/eurpub/ckq160

*Konkolÿ Thege, B., Bachner, Y. G., Kushnir, T., & Kopp, M. S. (2009). Relationship between meaning in life and smoking status: Results of a National Representative Survey. *Addictive Behaviors, 34,* 117–120. http://dx.doi.org/10.1016/j.addbeh.2008.09.001

*Kopp, M. S., Skrabski, A., & Szedmák, S. (1995). Socioeconomic factors, severity of depressive symptomatology, and sickness absence rate in the Hungarian population. *Journal of Psychosomatic Research, 39,* 1019–1029. http://dx.doi.org/10.1016/0022-3999(95)00513-7

Koss, M. P., Goodman, L. A., Browne, A., Fitzgerald, L. F., Keita, G. P., & Russo, N. F. (1994). *No safe haven: Male violence against women at home, at work, and in the community.* Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10156-000

Kovacs, M. (1985). The Children's Depression, Inventory (CDI). *Psychopharmacology Bulletin, 21,* 995–998.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16,* 606–613. http://dx.doi.org/10.1046/j.1525-1497.2001.016009606.x

Kuehner, C. (2003). Gender differences in unipolar depression: An update of epidemiological findings and possible explanations. *Acta Psychiatrica Scandinavica, 108,* 163–174. http://dx.doi.org/10.1034/j.1600-0447.2003.00204.x

**Lara, E., Olaya, B., Garin, N., Ayuso-Mateos, J. L., Miret, M., Moneta, V., & Haro, J. M. (2015). Is cognitive impairment associated with suicidality? A population-based study. *European Neuropsychopharmacology, 25,* 203–213. http://dx.doi.org/10.1016/j.euroneuro.2014.08.010

*Lee, M. A. (2015). Emotional abuse in childhood and suicidality: The mediating roles of re-victimization and depressive symptoms in adulthood. *Child Abuse & Neglect, 44,* 130–139. http://dx.doi.org/10.1016/j.chiabu.2015.03.016

*Lei, X., Sun, X., Strauss, J., Zhang, P., & Zhao, Y. (2014). Depressive symptoms and SES among the mid-aged and elderly in China: Evidence from the China Health and Retirement Longitudinal Study national baseline. *Social Science & Medicine, 120,* 224–232. http://dx.doi.org/10.1016/j.socscimed.2014.09.028

**Lépine, J. P., Gastpar, M., Mendlewicz, J., & Tylee, A. (1997). Depression in the community: The first pan-European study DEPRES (Depression Research in European Society). *International Clinical Psychopharmacology, 12,* 19–30. http://dx.doi.org/10.1097/00004850-199701000-00003

Lewinsohn, P. M., Rohde, P., Seeley, J. R., Klein, D. N., & Gotlib, I. H. (2000). Natural course of adolescent major depressive disorder in a community sample: Predictors of recurrence in young adults. *The American Journal of Psychiatry, 157,* 1584–1591. http://dx.doi.org/10.1176/appi.ajp.157.10.1584

Lewinsohn, P. M., Solomon, A., Seeley, J. R., & Zeiss, A. (2000). Clinical implications of "subthreshold" depressive symptoms. *Journal of Abnormal Psychology, 109,* 345–351. http://dx.doi.org/10.1037/0021-843X.109.2.345

Lewis, G., Pelosi, A. J., Araya, R., & Dunn, G. (1992). Measuring psychiatric disorder in the community: A standardized assessment for use by lay interviewers. *Psychological Medicine, 22,* 465–486. http://dx.doi.org/10.1017/S0033291700030415

Lindberg, S. M., Grabe, S., & Hyde, J. S. (2007). Gender, pubertal development, and peer sexual harassment predict objectified body consciousness in early adolescence. *Journal of Research on Adolescence, 17,* 723–742. http://dx.doi.org/10.1111/j.1532-7795.2007.00544.x

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136,* 1123–1135. http://dx.doi.org/10.1037/a0021276

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

**Lorenzo-Blanco, E. I., & Cortina, L. M. (2013). Latino/a depression and smoking: An analysis through the lenses of culture, gender, and ethnicity. *American Journal of Community Psychology, 51,* 332–346. http://dx.doi.org/10.1007/s10464-012-9553-3

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales*. Sydney, South Wales: Psychology Foundation.

Luppa, M., Sikorski, C., Luck, T., Ehreke, L., Konnopka, A., Wiese, B., . . . Riedel-Heller, S. G. (2012). Age- and gender-specific prevalence of depression in latest-life—Systematic review and meta-analysis. *Journal of Affective Disorders, 136,* 212–221. http://dx.doi.org/10.1016/j.jad.2010.11.033

*Madianos, M. G., & Stefanis, C. N. (1992). Changes in the prevalence of symptoms of depression and depression across Greece. *Social Psychiatry and Psychiatric Epidemiology, 27,* 211–219. http://dx.doi.org/10.1007/BF00788931

*Mäntyselkä, P. T., Turunen, J. H., Ahonen, R. S., & Kumpusalo, E. A. (2003). Chronic pain and poor self-rated health. *Journal of the American Medical Association, 290,* 2435–2442. http://dx.doi.org/10.1001/jama.290.18.2435

*Margraf, J., Lavallee, K., Zhang, X., & Schneider, S. (2016). Social rhythm and mental health: A cross-cultural comparison. *PLoS ONE, 11,* e0150312. http://dx.doi.org/10.1371/journal.pone.0150312

Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin, 137,* 856–879. http://dx.doi.org/10.1037/a0023678

*Marmorstein, N. R. (2009). Longitudinal associations between alcohol problems and depressive symptoms: Early adolescence through early adulthood. *Alcoholism: Clinical and Experimental Research, 33,* 49–59. http://dx.doi.org/10.1111/j.1530-0277.2008.00810.x

*,**Maske, U. E., Buttery, A. K., Beesdo-Baum, K., Riedel-Heller, S., Hapke, U., & Busch, M. A. (2016). Prevalence and correlates of *DSM–IV–TR* major depressive disorder, self-reported diagnosed depression and current depressive symptoms among adults in Germany. *Journal of Affective Disorders, 190,* 167–177. http://dx.doi.org/10.1016/j.jad.2015.10.006

**McMartin, S. E., Jacka, F. N., & Colman, I. (2013). The association between fruit and vegetable consumption and mental health disorders: Evidence from five waves of a national survey of Canadians. *Preventive Medicine, 56,* 225–230. http://dx.doi.org/10.1016/j.ypmed.2012.12.016

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science, 11,* 730–749. http://dx.doi.org/10.1177/1745691616662243

Mercier, R. J., Garrett, J., Thorp, J., & Siega-Riz, A. M. (2013). Pregnancy intention and postpartum depression: Secondary data analysis from a prospective cohort. *BJOG, 120,* 1116–1122. http://dx.doi.org/10.1111/1471-0528.12255

**Merikangas, A. K., Mendola, P., Pastor, P. N., Reuben, C. A., & Cleary, S. D. (2012). The association between major depressive disorder and obesity in US adolescents: Results from the 2001–2004 National Health and Nutrition Examination Survey. *Journal of Behavioral Medicine, 35,* 149–154. http://dx.doi.org/10.1007/s10865-011-9340-x

*Michal, M., Glaesmer, H., Zwerenz, R., Knebel, A., Wiltink, J., Brähler, E., & Beutel, M. E. (2011). Base rates for depersonalization according to the 2-item version of the Cambridge Depersonalization Scale (CDS-2) and its associations with depression/anxiety in the general population. *Journal of Affective Disorders, 128,* 106–111. http://dx.doi.org/10.1016/j.jad.2010.06.033

Mirowsky, J. (1996). Age and the gender gap in depression. *Journal of Health and Social Behavior, 37,* 362–380. http://dx.doi.org/10.2307/2137263

**Mohammadi, M. R., Davidian, H., Noorbala, A. A., Malekafzali, H., Naghavi, H. R., Pouretemad, H. R., . . . Ghanizadeh, A. (2005). An epidemiological survey of psychiatric disorders in Iran. *Clinical Practice and Epidemiology in Mental Health, 1,* 16. http://dx.doi.org/10.1186/1745-0179-1-16

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6,* e1000097. http://dx.doi.org/10.1371/journal.pmed.1000097

**Mommersteeg, P. M., Herr, R., Pouwer, F., Holt, R. I., & Loerbroks, A. (2013). The association between diabetes and an episode of depressive symptoms in the 2002 World Health Survey: An analysis of 231,797 individuals from 47 countries. *Diabetic Medicine, 30,* e208–e214. http://dx.doi.org/10.1111/dme.12193

*Momtaz, Y. A., Hamid, T. A., Haron, S. A., & Bagat, M. F. (2016). Flourishing in later life. *Archives of Gerontology and Geriatrics, 63,* 85–91. http://dx.doi.org/10.1016/j.archger.2015.11.001

Moos, R. H., Cronkite, R. C., Billings, A. G., & Finney, J. W. (1985). *Health and Daily Living Form manual, revised version*. Stanford, CA: Social Ecological Laboratory, Veterans Administration and Stanford University Medical Centers.

*Morozink, J. A., Friedman, E. M., Coe, C. L., & Ryff, C. D. (2010). Socioeconomic and psychosocial predictors of interleukin-6 in the MIDUS national sample. *Health Psychology, 29,* 626–635. http://dx.doi.org/10.1037/a0021360

*Mumford, E. A., Liu, W., Hair, E. C., & Yu, T. C. (2013). Concurrent trajectories of BMI and mental health patterns in emerging adulthood. *Social Science & Medicine, 98,* 1–7. http://dx.doi.org/10.1016/j.socscimed.2013.08.036

*Munhoz, T. N., Nunes, B. P., Wehrmeister, F. C., Santos, I. S., & Matijasevich, A. (2016). A nationwide population-based study of depression in Brazil. *Journal of Affective Disorders, 192,* 226–233. http://dx.doi.org/10.1016/j.jad.2015.12.038

Muñoz, R. F., Cuijpers, P., Smit, F., Barrera, A. Z., & Leykin, Y. (2010). Prevention of major depression. *Annual Review of Clinical Psychology, 6,* 181–212. http://dx.doi.org/10.1146/annurev-clinpsy-033109-132040

National Research Institute of Public Health, Russian Academy of Medical Sciences (RAMS). (2013). *The Russian Federation: Study on global AGEing and adult health (SAGE), Wave 1*. Geneva, Switzerland: World Health Organization.

*Neumark-Sztainer, D., & Hannan, P. J. (2000). Weight-related behaviors among adolescent girls and boys: Results from a national survey. *Archives of Pediatrics & Adolescent Medicine, 154,* 569–577. http://dx.doi.org/10.1001/archpedi.154.6.569

Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin, 101,* 259–282. http://dx.doi.org/10.1037/0033-2909.101.2.259

O'Connor, E. A., Whitlock, E. P., Beil, T. L., & Gaynes, B. N. (2009). Screening for depression in adult patients in primary care settings: A systematic evidence review. *Annals of Internal Medicine, 151,* 793–803. http://dx.doi.org/10.7326/0003-4819-151-11-200912010-00007

*Oh, D. H., Kim, S. A., Lee, H. Y., Seo, J. Y., Choi, B. Y., & Nam, J. H. (2013). Prevalence and correlates of depressive symptoms in Korean adults: Results of a 2009 korean community health survey. *Journal of Korean Medical Science, 28,* 128–135. http://dx.doi.org/10.3346/jkms.2013.28.1.128

*O'Halloran, A. M., Kenny, R. A., & King-Kallimanis, B. L. (2014). The latent factors of depression from the short forms of the CES-D are consistent, reliable and valid in community-living older adults. *European Geriatric Medicine, 5,* 97–102. http://dx.doi.org/10.1016/j.eurger.2013.12.004

*Ojard, C., Donnelly, J. P., Safford, M. M., Griffin, R., & Wang, H. E. (2015). Psychosocial stress as a risk factor for sepsis: A population-based cohort study. *Psychosomatic Medicine, 77,* 93–100. http://dx.doi.org/10.1097/PSY.0000000000000120

*Okabayashi, H., Liang, J., Krause, N., Akiyama, H., & Sugisawa, H. (2004). Mental health among older adults in Japan: Do sources of social support and negative interaction make a difference? *Social Science & Medicine, 59,* 2259–2270. http://dx.doi.org/10.1016/j.socscimed.2004.02.024

*Oksuzyan, A., Crimmins, E., Saito, Y., O'Rand, A., Vaupel, J. W., & Christensen, K. (2010). Cross-national comparison of sex differences in health and mortality in Denmark, Japan and the US. *European Journal of Epidemiology, 25,* 471–480. http://dx.doi.org/10.1007/s10654-010-9460-6

*Olsen, L. R., Mortensen, E. L., & Bech, P. (2006). Mental distress in the Danish general population. *Acta Psychiatrica Scandinavica, 113,* 477–484. http://dx.doi.org/10.1111/j.1600-0447.2005.00743.x

Oquendo, M. A., Ellis, S. P., Greenwald, S., Malone, K. M., Weissman, M. M., & Mann, J. J. (2001). Ethnic and sex differences in suicide rates relative to major depression in the United States. *The American Journal of Psychiatry, 158,* 1652–1658. http://dx.doi.org/10.1176/appi.ajp.158.10.1652

*Ostbye, T., Malhotra, R., & Chan, A. (2009). Thirteen dimensions of health in elderly Sri Lankans: Results from a National Sri Lanka Aging Survey. *Journal of the American Geriatrics Society, 57,* 1376–1387. http://dx.doi.org/10.1111/j.1532-5415.2009.02350.x

*Park, J. H., Kim, K. W., Kim, M. H., Kim, M. D., Kim, B. J., Kim, S. K., . . . Cho, M. J. (2012). A nationwide survey on the prevalence and risk factors of late life depression in South Korea. *Journal of Affective Disorders, 138*(1–2), 34–40. http://dx.doi.org/10.1016/j.jad.2011.12.038

**Patten, S. B. (2005). Markov models of major depression for linking psychiatric epidemiology to clinical practice. *Clinical Practice and Epidemiology in Mental Health, 1,* 2. http://dx.doi.org/10.1186/1745-0179-1-2

Patten, S. B., Williams, J. V., Lavorato, D. H., Wang, J. L., Bulloch, A. G., & Sajobi, T. (2016). The association between major depression prevalence and sex becomes weaker with age. *Social Psychiatry and Psychiatric Epidemiology, 51,* 203–210. http://dx.doi.org/10.1007/s00127-015-1166-3

Pattyn, E., Verhaeghe, M., & Bracke, P. (2015). The gender gap in mental health service use. *Social Psychiatry and Psychiatric Epidemiology, 50,* 1089–1095. http://dx.doi.org/10.1007/s00127-015-1038-x

**Peltzer, K., & Phaswana-Mafuya, N. (2013). Depression and associated factors in older adults in South Africa. *Global Health Action, 6,* 1–9. http://dx.doi.org/10.3402/gha.v6i0.18871

Petersen, J. L., & Hyde, J. S. (2010). A meta-analytic review of research on gender differences in sexuality, 1993–2007. *Psychological Bulletin, 136,* 21–38. http://dx.doi.org/10.1037/a0017504

Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression. Critical review. *The British Journal of Psychiatry, 177,* 486–492. http://dx.doi.org/10.1192/bjp.177.6.486

**Pirkola, S. P., Isometsä, E., Suvisaari, J., Aro, H., Joukamaa, M., Poikolainen, K., . . . Lönnqvist, J. K. (2005). DSM–IV mood-, anxiety- and alcohol use disorders and their comorbidity in the Finnish general population—Results from the Health 2000 Study. *Social Psychiatry and Psychiatric Epidemiology, 40,* 1–10. http://dx.doi.org/10.1007/s00127-005-0848-7

Prince, M. J., Reischies, F., Beekman, A. T. F., Fuhrer, R., Jonker, C., Kivela, S. L., . . . Copeland, J. R. (1999). Development of the EURO-D Scale—A European, Union initiative to compare symptoms of depression in 14 European centres. *The British Journal of Psychiatry, 174,* 330–338. http://dx.doi.org/10.1192/bjp.174.4.330

Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for research in the general population. *Applied Psychological Measurement, 1,* 385–401. http://dx.doi.org/10.1177/014662167700100306

**Rafful, C., Medina-Mora, M. E., Borges, G., Benjet, C., & Orozco, R. (2012). Depression, gender, and the treatment gap in Mexico. *Journal of Affective Disorders, 138,* 165–169. http://dx.doi.org/10.1016/j.jad.2011.12.040

**Rapsey, C. M., Lim, C. C., Al-Hamzawi, A., Alonso, J., Bruffaerts, R., Caldas-de-Almeida, J. M., . . . Scott, K. M. (2015). Associations between *DSM–IV* mental disorders and subsequent COPD diagnosis. *Journal of Psychosomatic Research, 79,* 333–339. http://dx.doi.org/10.1016/j.jpsychores.2015.08.005

Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of Educational Progress Assessments. *Journal of Educational Psychology, 107,* 645–662. http://dx.doi.org/10.1037/edu0000012

Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Social Science & Medicine, 90,* 24–31. http://dx.doi.org/10.1016/j.socscimed.2013.04.026

Revah-Levy, A., Birmaher, B., Gasquet, I., & Falissard, B. (2007). The Adolescent Depression Rating Scale (ADRS): A validation study. *BMC Psychiatry, 7,* 2. http://dx.doi.org/10.1186/1471-244X-7-2

*Revah-Levy, A., Speranza, M., Barry, C., Hassler, C., Gasquet, I., Moro, M. R., & Falissard, B. (2011). Association between Body Mass Index and depression: The "fat and jolly" hypothesis for adolescents girls. *BMC Public Health, 11,* 649. http://dx.doi.org/10.1186/1471-2458-11-649

*Rey, J. M., Sawyer, M. G., Clark, J. J., & Baghurst, P. A. (2001). Depression among Australian adolescents. *The Medical Journal of Australia, 175,* 19–23.

Reynolds, W. M. (1986). *RADS professional manual.* Odessa, FL: Psychological Assessment Resources.

*Rief, W., Glaesmer, H., Baehr, V., Broadbent, E., Brähler, E., & Petrie, K. J. (2012). The relationship of modern health worries to depression, symptom reporting and quality of life in a general population survey. *Journal of Psychosomatic Research, 72,* 318–320. http://dx.doi.org/10.1016/j.jpsychores.2011.11.017

Rindfuss, R. R., Swicegood, C. G., & Rosenfeld, R. A. (1987). Disorder in the life course: How common and does it matter? *American Sociological Review, 52,* 785–801. http://dx.doi.org/10.2307/2095835

*Risal, A., Manandhar, K., Linde, M., Steiner, T. J., & Holen, A. (2016). Anxiety and depression in Nepal: Prevalence, comorbidity and associations. *BMC Psychiatry, 16,* 102. http://dx.doi.org/10.1186/s12888-016-0810-0

Robins, L. N., Helzer, J. E., Croughan, J., & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule. Its history, characteristics, and validity. *Archives of General Psychiatry, 38,* 381–389. http://dx.doi.org/10.1001/archpsyc.1981.01780290015001

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641. http://dx.doi.org/10.1037/0033-2909.86.3.638

Salk, R. H., Petersen, J. L., Abramson, L. Y., & Hyde, J. S. (2016). The contemporary face of gender differences and similarities in depression throughout adolescence: Development and chronicity. *Journal of Affective Disorders, 205,* 28–35. http://dx.doi.org/10.1016/j.jad.2016.03.071

Salokangas, R. K., Poutanen, O., & Stengård, E. (1995). Screening for depression in primary care. Development and validation of the Depression Scale, a screening instrument for depression. *Acta Psychiatrica Scandinavica, 92,* 10–16. http://dx.doi.org/10.1111/j.1600-0447.1995.tb09536.x

**Sandanger, I., Nygård, J. F., Sørensen, T., & Dalgard, O. S. (2007). Return of depressed men: Changes in distribution of depression and symptom cases in Norway between 1990 and 2001. *Journal of Affective Disorders, 100,* 153–162. http://dx.doi.org/10.1016/j.jad.2006.10.006

*Sandman, N., Valli, K., Kronholm, E., Revonsuo, A., Laatikainen, T., & Paunio, T. (2015). Nightmares: Risk factors among the Finnish general adult population. *Sleep, 38,* 507–514.

*Scafato, E., Galluzzo, L., Ghirini, S., Gandin, C., Rossi, A., Solfrizzi, V., . . . the ILSA Working Group. (2012). Changes in severity of depressive symptoms and mortality: The Italian Longitudinal Study on Aging. *Psychological Medicine, 42,* 2619–2629. http://dx.doi.org/10.1017/S0033291712000645

**Scott, K. M., Oakley Browne, M. A., & Elisabeth Wells, J. (2010). Prevalence, impairment and severity of 12-month *DSM–IV* major depressive episodes in Te Rau Hinengaro: New Zealand Mental Health Survey 2003/4. *Australian and New Zealand Journal of Psychiatry, 44,* 750–758. http://dx.doi.org/10.3109/00048671003781780

*Seaton, E. K., Caldwell, C. H., Sellers, R. M., & Jackson, J. S. (2008). The prevalence of perceived discrimination among African American and Caribbean Black youth. *Developmental Psychology, 44,* 1288–1297. http://dx.doi.org/10.1037/a0012747

Seedat, S., Scott, K. M., Angermeyer, M. C., Berglund, P., Bromet, E. J., Brugha, T. S., . . . Kessler, R. C. (2009). Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. *Archives of General Psychiatry, 66,* 785–795. http://dx.doi.org/10.1001/archgenpsychiatry.2009.36

*Seppälä, J., Koponen, H., Kautiainen, H., Eriksson, J. G., Kampman, O., Männistö, S., . . . Vanhala, M. (2012). Association between folate intake and melancholic depressive symptoms. A Finnish population-based study. *Journal of Affective Disorders, 138,* 473–478. http://dx.doi.org/10.1016/j.jad.2012.01.007

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry, 39,* 28–38. http://dx.doi.org/10.1097/00004583-200001000-00014

**Shah, A. J., Veledar, E., Hong, Y., Bremner, J. D., & Vaccarino, V. (2011). Depression and history of attempted suicide as risk factors for heart disease mortality in young individuals. *Archives of General Psychiatry, 68,* 1135–1142. http://dx.doi.org/10.1001/archgenpsychiatry.2011.125

Shanghai Municipal Center for Disease Control and Prevention (SCDC). (2012). *China: Study on global AGEing and adult health (SAGE), Wave 1.* Geneva, Switzerland: World Health Organization.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M. I. N. I.): The development and validation of a structured diagnostic psychiatric interview for *DSM–IV* and *ICD-10*. *The Journal of Clinical Psychiatry, 59*(Suppl. 20), 22–33.

*Shiovitz-Ezra, S., Leitsch, S., Graber, J., & Karraker, A. (2009). Quality of life and psychological health indicators in the national social life, health, and aging project. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 64*(Suppl. 1), i30–i37. http://dx.doi.org/10.1093/geronb/gbn020

Siegel, J. M., Aneshensel, C. A., Taub, B., Cantwell, D. P., & Driscoll, A. K. (1998). Adolescent depressed mood in a multiethnic sample. *Journal of Youth and Adolescence, 27,* 413–427. http://dx.doi.org/10.1023/A:1022873601030

*Sigfusdottir, I. D., Asgeirsdottir, B. B., Sigurdsson, J. F., & Gudjonsson, G. H. (2008). Trends in depressive symptoms, anxiety symptoms and visits to healthcare specialists: A national study among Icelandic adolescents. *Scandinavian Journal of Public Health, 36,* 361–368. http://dx.doi.org/10.1177/1403494807088457

**Skapinakis, P., Bellos, S., Koupidis, S., Grammatikopoulos, I., Theodorakis, P. N., & Mavreas, V. (2013). Prevalence and sociodemographic associations of common mental disorders in a nationally representative sample of the general population of Greece. *BMC Psychiatry, 13,* 163. http://dx.doi.org/10.1186/1471-244X-13-163

**Slade, T., Johnston, A., Oakley Browne, M. A., Andrews, G., & Whiteford, H. (2009). 2007 National Survey of Mental Health and Wellbeing: Methods and key findings. *The Australian and New Zealand Journal of Psychiatry, 43,* 594–605. http://dx.doi.org/10.1080/00048670902970882

Smeets, R. M. W., & Dingemans, P. M. A. J. (1993). *Composite International Diagnostic Interview (CIDI), version 1.1.* Amsterdam, Geneva: World Health Organization.

*Song, L. (2011). Social capital and psychological distress. *Journal of Health and Social Behavior, 52,* 478–492. http://dx.doi.org/10.1177/0022146511411921

*Sonnenberg, C. M., Deeg, D. J., van Tilburg, T. G., Vink, D., Stek, M. L., & Beekman, A. T. (2013). Gender differences in the relation between depression and social support in later life. *International Psychogeriatrics, 25,* 61–70. http://dx.doi.org/10.1017/S1041610212001202

**Spiers, N., Brugha, T. S., Bebbington, P., McManus, S., Jenkins, R., & Meltzer, H. (2012). Age and birth cohort differences in depression in repeated cross-sectional surveys in England: The National Psychiatric Morbidity Surveys, 1993 to 2007. *Psychological Medicine, 42,* 2047–2055. http://dx.doi.org/10.1017/S003329171200013X

*Steptoe, A., Demakakos, P., de Oliveira, C., & Wardle, J. (2012). Distinctive biological correlates of positive psychological well-being in older men and women. *Psychosomatic Medicine, 74,* 501–508. http://dx.doi.org/10.1097/PSY.0b013e31824f82c8

*Strong, C., Tsai, M.-C., Lin, C-Y., & Cheng, C.-P. (2016). Childhood adversity, timing of puberty and adolescent depressive symptoms: A longitudinal study in Taiwan. *Child Psychiatry and Human Development, 47,* 347–357. http://dx.doi.org/10.1007/s10578-015-0570-y

**Subramaniam, M., Abdin, E., Vaingankar, J. A., Nan, L., Heng, D., McCrone, P., & Chong, S. A. (2013). Impact of psychiatric disorders and chronic physical conditions on health-related quality of life: Singapore Mental Health Study. *Journal of Affective Disorders, 147,* 325–330. http://dx.doi.org/10.1016/j.jad.2012.11.033

*Sugihara, Y., Sugisawa, H., Shibata, H., & Harada, K. (2008). Productive roles, gender, and depressive symptoms: Evidence from a national longitudinal study of late-middle-aged Japanese. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences, 63,* P227–P234. http://dx.doi.org/10.1093/geronb/63.4.P227

**Suttajit, S., Kittirattanapaiboon, P., Junsirimongkol, B., Likhitsathian, S., & Srisurapanont, M. (2012). Risks of major depressive disorder and anxiety disorders among Thais with alcohol use disorders and illicit drug use: Findings from the 2008 Thai National Mental Health Survey. *Addictive Behaviors, 37,* 1395–1399. http://dx.doi.org/10.1016/j.addbeh.2012.06.014

*Symonds, J., Dietrich, J., Chow, A., & Salmela-Aro, K. (2016). Mental health improves after transition from comprehensive school to vocational education or employment in England: A national cohort study. *Developmental Psychology, 52,* 652–665. http://dx.doi.org/10.1037/a0040118

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York, NY: HarperCollins.

**Takeuchi, D. T., Zane, N., Hong, S., Chae, D. H., Gong, F., Gee, G. C., . . . Alegría, M. (2007). Immigration-related factors and mental disorders

among Asian Americans. *American Journal of Public Health, 97,* 84–90. http://dx.doi.org/10.2105/AJPH.2006.088401

*Thege, B. K., Bachner, Y. G., Kushnir, T., & Kopp, M. S. (2009). Relationship between meaning in life and smoking status: Results of a national representative survey. *Addictive Behaviors, 34,* 117–120. http://dx.doi.org/10.1016/j.addbeh.2008.09.001

*Thibodeau, M. A., & Asmundson, G. J. G. (2014). The PHQ-9 assesses depression similarly in men and women from the general population. *Personality and Individual Differences, 56,* 149–153. http://dx.doi.org/10.1016/j.paid.2013.08.039

*Torikka, A., Kaltiala-Heino, R., Rimpelä, A., Marttunen, M., Luukkaala, T., & Rimpelä, M. (2014). Self-reported depression is increasing among socio-economically disadvantaged adolescents - repeated cross-sectional surveys from Finland from 2000 to 2011. *BMC Public Health, 14,* 408. http://dx.doi.org/10.1186/1471-2458-14-408

*Torres, J. M., & Wong, R. (2013). Childhood poverty and depressive symptoms for older adults in Mexico: A life-course analysis. *Journal of Cross-Cultural Gerontology, 28,* 317–337. http://dx.doi.org/10.1007/s10823-013-9198-1

**Toussaint, L. L., Williams, D. R., Musick, M. A., & Everson-Rose, S. A. (2008). The association of forgiveness and 12-month prevalence of major depressive episode: Gender differences in a probability sample of U.S. adults. *Mental Health, Religion & Culture, 11,* 485–500. http://dx.doi.org/10.1080/13674670701564989

Tsai, J. L., & Chentsova-Dutton, Y. (2002). Understanding depression across cultures. In I. H. Gotlib & C. L. Hammen (Eds.), *Handbook of depression* (pp. 467–491). New York, NY: Guilford Press.

Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, socioeconomic status, and birth cohort differences on the Children's Depression Inventory: A meta-analysis. *Journal of Abnormal Psychology, 111,* 578–588. http://dx.doi.org/10.1037/0021-843X.111.4.578

United Nations. (2010). *The world's women*. Retrieved from http://unstats.un.org/unsd/demographic/products/Worldswomen/WW_full%20report_color.pdf

United Nations Development Programme. (2005). *Human development report 2005*. Retrieved from http://hdr.undp.org/en/content/human-development-report-2005

University of Ghana Medical School, Department of Community Health. (2013). *Ghana: Study on global AGEing and adult health (SAGE), Wave 1*. Geneva, Switzerland: World Health Organization.

*Van de Velde, S., Bracke, P., & Levecque, K. (2010). Gender differences in depression in 23 European countries. Cross-national variation in the gender gap in depression. *Social Science & Medicine, 71,* 305–313. http://dx.doi.org/10.1016/j.socscimed.2010.03.035

Van de Velde, S., Huijts, T., Bracke, P., & Bambra, C. (2013). Macro-level gender equality and depression in men and women in Europe. *Sociology of Health & Illness, 35,* 682–698. http://dx.doi.org/10.1111/j.1467-9566.2012.01521.x

*van Praag, L., Bracke, P., Christiaens, W., Levecque, K., & Pattyn, E. (2009). Mental health in a gendered context: Gendered community effect on depression and problem drinking. *Health & Place, 15,* 990–998. http://dx.doi.org/10.1016/j.healthplace.2009.04.003

Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods, 10,* 428–443. http://dx.doi.org/10.1037/1082-989X.10.4.428

**Vicente, B., Kohn, R., Rioseco, P., Saldivia, S., Levav, I., & Torres, S. (2006). Lifetime and 12-month prevalence of *DSM–III–R* disorders in the Chile psychiatric prevalence study. *The American Journal of Psychiatry, 163,* 1362–1370. http://dx.doi.org/10.1176/ajp.2006.163.8.1362

*Villatoro, J. A., Medina-Mora, M. E., Juárez, F., Rojas, E., Carreño, S., & Berenzon, S. (1998). Drug use pathways among high school students of Mexico. *Addiction, 93,* 1577–1588. http://dx.doi.org/10.1046/j.1360-0443.1998.9310157715.x

**Volken, T. (2013). Second-stage non-response in the Swiss health survey: Determinants and bias in outcomes. *BMC Public Health, 13,* 167. http://dx.doi.org/10.1186/1471-2458-13-167

*von Soest, T., & Wichstrøm, L. (2014). Secular trends in depressive symptoms among Norwegian adolescents from 1992 to 2010. *Journal of Abnormal Child Psychology, 42,* 403–415. http://dx.doi.org/10.1007/s10802-013-9785-1

**Wade, T. J., Cairney, J., & Pevalin, D. J. (2002). Emergence of gender differences in depression during adolescence: National panel results from three countries. *Journal of the American Academy of Child & Adolescent Psychiatry, 41,* 190–198. http://dx.doi.org/10.1097/00004583-200202000-00013

*Walker, L., Merry, S., Watson, P. D., Robinson, E., Crengle, S., & Schaaf, D. (2005). The Reynolds Adolescent Depression Scale in New Zealand adolescents. *Australian and New Zealand Journal of Psychiatry, 39,* 136–140. http://dx.doi.org/10.1080/j.1440-1614.2005.01534.x

*Wang, J., Iannotti, R. J., Luk, J. W., & Nansel, T. R. (2010). Co-occurrence of victimization from five subtypes of bullying: Physical, verbal, social exclusion, spreading rumors, and cyber. *Journal of Pediatric Psychology, 35,* 1103–1112. http://dx.doi.org/10.1093/jpepsy/jsq048

Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care, 30,* 473–483. http://dx.doi.org/10.1097/00005650-199206000-00002

Weissman, M. M., & Klerman, G. L. (1977). Sex differences and the epidemiology of depression. *Archives of General Psychiatry, 34,* 98–111. http://dx.doi.org/10.1001/archpsyc.1977.01770130100011

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

Wilson, D. B. (2006). *Meta-analysis macros for SAS, SPSS, and Stata*. Retrieved from http://mason.gmu.edu/~dwilsonb/ma.html

Wilson, D. B. (2015). *Practical meta-analysis effect size calculator* [Online calculator]. Retrieved from https://www.campbellcollaboration.org/this-is-a-web-based-effect-size-calculator/explore/this-is-a-web-based-effect-size-calculator

**Wittchen, H. U., Carter, R. M., Pfister, H., Montgomery, S. A., & Kessler, R. C. (2000). Disabilities and quality of life in pure and comorbid generalized anxiety disorder and major depression in a national survey. *International Clinical Psychopharmacology, 15,* 319–328. http://dx.doi.org/10.1097/00004850-200015060-00002

Wittchen, H. U., & Pfister, H. (1997). *DIA-X-Interviews: Manual für Screening-Verfahren und Interview; Interviewheft Längsschnittuntersuchung (DIA-X-Lifetime); Er- gänzungsheft (DIA-X-Lifetime); Interviewheft Querschnittuntersuchung (DIA- X-12Monate); Ergänzungsheft (DIA-X-12Monate); PC-Programm zur Durchführung des Interviews (Längsund Querschnittuntersuchung); Auswertungsprogramm*. Frankfurt, Germany: Swets & Zeitlinger.

Wittchen, H. U., Robins, L. N., Cottler, L. B., Sartorius, N., Burke, J. D., & Regier, D. (1991). Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview. *The British Journal of Psychiatry, 159,* 645–653, 658. http://dx.doi.org/10.1192/bjp.159.5.645

Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. *Advances in Experimental Social Psychology, 46,* 55–123. http://dx.doi.org/10.1016/B978-0-12-394281-4.00002-7

World Bank. (2004). *GINI index*. Retrieved from http://data.worldbank.org/indicator/SI.POV.GINI?page=2&order=wbapi_data_value_2011%20wbapi_data_value%20wbapi_data_value-first&sort=asc

World Health Organization. (1990). *Composite International Diagnostic Interview (CIDI), version 1.0*. Geneva, Switzerland: Author.

World Health Organization. (1992). *The ICD-10* classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines. Geneva, Switzerland: Author.

World Health Organization. (1997). *Composite International Diagnostic Interview, Version 2.1.* Geneva, Switzerland: Author.

World Health Organization. (2002). *World Health Survey.* Retrieved from http://www.who.int/healthinfo/survey/en/index.html

World Health Organization. (2016). *Out of the shadows: Making mental health a global development priority.* Retrieved from http://www.who.int/mental_health/advocacy/wb_background_paper.pdf?ua=1

World Values Survey. (2014). *Wave 4, 1999–2004.* Retrieved from http://www.worldvaluessurvey.org/WVSDocumentationWV4.jsp

*Yamada, K., & Teerawichitchainan, B. (2015). Living arrangements and psychological well-being of the older adults after the economic transition in Vietnam. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences, 70,* 957–968. http://dx.doi.org/10.1093/geronb/gbv059

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a Geriatric Depression Screening Scale: A preliminary report. *Journal of Psychiatric Research, 17,* 37–49. http://dx.doi.org/10.1016/0022-3956(82)90033-4

*Zemore, S. E., Karriker-Jaffe, K. J., & Mulia, N. (2013). Temporal trends and changing racial/ethnic disparities in alcohol problems: Results from the 2000 to 2010 National Alcohol Surveys. *Journal of Addiction Research & Therapy.* Advance online publication.

Zentner, M., & Mitura, K. (2012). Stepping out of the caveman's shadow: Nations' gender gap predicts degree of sex differentiation in mate preferences. *Psychological Science, 23,* 1176–1185. http://dx.doi.org/10.1177/0956797612441004

**Zhao, W., Chen, Y., Lin, M., & Sigal, R. J. (2006). Association between diabetes and depression: Sex and age differences. *Public Health, 120,* 696–704. http://dx.doi.org/10.1016/j.puhe.2006.04.012

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67,* 361–370. http://dx.doi.org/10.1111/j.1600-0447.1983.tb09716.x

**Zinzow, H. M., Ruggiero, K. J., Resnick, H., Hanson, R., Smith, D., Saunders, B., & Kilpatrick, D. (2009). Prevalence and mental health correlates of witnessed parental and community violence in a national sample of adolescents. *Journal of Child Psychology and Psychiatry, 50,* 441–450. http://dx.doi.org/10.1111/j.1469-7610.2008.02004.x

**Zubrick, S. R., Hafekost, J., Johnson, S. E., Lawrence, D., Saw, S., Sawyer, M., . . . Buckingham, W. J. (2016). Suicidal behaviours: Prevalence estimates from the second Australian Child and Adolescent Survey of Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry, 50,* 899–910. http://dx.doi.org/10.1177/0004867415622563

*Zunzunegui, M. V., Minicuci, N., Blumstein, T., Noale, M., Deeg, D., Jylhä, M., . . . the CLESA Working Group. (2007). Gender differences in depressive symptoms among older adults: A cross-national comparison: The CLESA project. *Social Psychiatry and Psychiatric Epidemiology, 42,* 198–207. http://dx.doi.org/10.1007/s00127-007-0158-3