Development of Depression From Preadolescence to Young Adulthood: Emerging Gender Differences in a 10-Year Longitudinal Study

Benjamin L. Hankin, Lyn Y. Abramson, and Terrie E. Moffitt
University of Wisconsin—Madison

Phil A. Silva and Rob McGee
University of Otago Medical School

Kathryn E. Angell
University of Wisconsin—Madison

The authors investigated the emergence of gender differences in clinical depression and the overall development of depression from preadolescence to young adulthood among members of a complete birth cohort using a prospective longitudinal approach with structured diagnostic interviews administered 5 times over the course of 10 years. Small gender differences in depression (females greater than males) first began to emerge between the ages of 13 and 15. However, the greatest increase in this gender difference occurred between ages 15 and 18. Depression rates and accompanying gender differences for a university student subsample were no different than for a nonuniversity subsample. There was no gender difference for depression recurrence or for depression symptom severity. The peak increase in both overall rates of depression and new cases of depression occurred between the ages of 15 and 18. Results suggest that middle-to-late adolescence (ages 15–18) may be a critical time for studying vulnerability to depression because of the higher depression rates and the greater risk for depression onset and dramatic increase in gender differences in depression during this period.

A consistent and robust finding across many countries is that approximately twice as many adult women are depressed as adult men (Nolen-Hoeksema, 1990; Weissman & Klerman, 1977; Weissman, Leaf, Holzer, Myers, & Tischler, 1984). On the other hand, preadolescent boys are more likely to be depressed than preadolescent girls in both clinic samples (Kashani, Cantwell, Shekim, & Reid, 1982) and in the general population (Anderson, Williams, McGee, & Silva, 1987; Rutter, 1977). Thus, sometime in adolescence a switch occurs such that at least twice as many girls are depressed as boys. Pinpointing exactly when this transition occurs may provide important information about causal mechanisms underlying the development of clinical depression. On the basis of the currently available data, recent reviewers (Nolen-Hoeksema, 1990; Petersen et al., 1993) have suggested that the switch in gender differences in depression emerges by age 14 to 15.

Temporal Resolution of the Emergence of Gender Differences in Depression

However, the empirical foundation for describing a timeline of the development of gender differences in depression is not as clear and consolidated as it may appear. First, although the gender difference in depression should emerge by age 14 to 15 given the above reviewers’ suggestion, previous studies examining this issue in adolescence have not found a Gender × Age interaction for ages 14–18 (Allgood-Merten, Lewinsohn, & Hops, 1990; Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993). Second, the two longitudinal studies (Ge, Lorenz, Conger, Elder, & Simons, 1994; Petersen, Sarigiani, & Kennedy, 1991) that have followed the same individuals from preadolescence to late adolescence did not assess the adolescents for a clinical diagnosis of depression. Of note, however, is that both studies did find that girls evidenced an increase in depressive affect (Petersen et al., 1991) and an increase in depressive symptoms (Ge et al., 1994) starting in the 8th grade (approximately age 13 to 14), whereas the depressive affect and symptoms for boys stayed relatively constant. These studies are important in understanding the progression of depressive mood and symptoms, yet they cannot provide a definitive answer to when the gender difference in clinically diagnosed depression emerges.

Overall, the main problem with the empirical literature to date is that no single study has longitudinally followed the same group of individuals in the community and recorded actual clinical diagnostic depression data at multiple ages to assess accurately the emergence of depression over time. Consequently, researchers have had to piece together the various cross-sectional epidemiological studies that have used different measures and different groups of people at different ages in an attempt to delineate the developmental progression of depression for males and females. Although combining cross-sectional studies is informative, this methodological strategy does not provide as compelling and clear a depiction of the emergence of gender...
differences in depression as would a longitudinal study. Furthermore, interpretations based on this methodological strategy may be confounded by documented cohort differences in rates of depression (e.g., Klerman et al., 1985; Myers et al., 1984; Robbins et al., 1984) and differential estimates of rates of depression yielded by different measures such as self-report questionnaire assessments of depressive symptoms compared with interview-based assessments of clinical diagnoses of depression (e.g., Coyne, 1994; Vredenburg, Flett, & Krames, 1993).

Delineating a precise developmental timeline for the emergence of depression separately by gender is important because theories about the correlates and causes of depression should be able to explain when these gender differences begin to emerge and how these differences progress over time. Thus, the major goal of this study was to use more powerful methods to draw such a developmental timeline for the emergence of depression by gender. We accomplish this by analyzing the diagnostic depression data from a prospective longitudinal study in Dunedin, New Zealand, which has collected this information in a population-based sample of the same individuals at multiple ages from age 11 to age 21. We trace both the overall rates of depression1 over the 10-year span as well as the new incidence rates (i.e., new cases) of depression. Results from these analyses should provide much clearer temporal resolution of the emergence of gender differences in depression. We predict that the gender difference in depression should first become apparent around early adolescence (age 13–14) because previous studies have found that the gender difference is already present in the 14–18 age range and because previous longitudinal studies observed an increase in girls’ depressive affect and symptoms around early adolescence.

Student Versus Nonstudent Samples

The second goal of this study was to examine an interesting anomaly in research on gender differences in depression: Gender differences in depression are not consistently exhibited among samples of university students. A brief review of studies focusing on depressive symptomatology (not on clinical diagnosis of depression) among university samples reveals this inconsistency. In some studies, gender differences were not found (Gladstone & Koenig, 1994; Hammen & Padesky, 1977), whereas in another study gender differences were found (Boggiano & Barrett, 1991) and, finally, in a third study, gender differences were or were not found depending on the self-report instrument used (Gotlib, 1984). Despite these clearly mixed results, it is widely assumed that university student groups are an exception to the general rule of gender differences in depression. Although no study to date has directly compared a university student group against a same-age nonstudent group (see Angell, Hankin, Abramson, Hyde, & Baker, 1995), it is generally assumed that a nonstudent group shows the usual 2:1 gender difference in depression because epidemiological studies (e.g., Myers et al., 1984) have demonstrated that such a gender difference does exist for the overall group of 18–24-year-old individuals, which includes both students and nonstudents. Resolving the controversy of whether university students demonstrate gender differences in depression is important not only because the true pattern of gender differences needs clarification, but also because the vast amount of psychological research on depression is conducted with university students (e.g., see Coyne, 1994; Vredenburg et al., 1993, for reviews). If university students do not exhibit as fundamental a feature of depression as higher rates for females than males, then the generalizability of this corpus of research with university undergraduates to the general population is questionable.

However, it is necessary to go beyond a simple cross-sectional comparison of a university student group versus a nonstudent group at one point in time. To date, the one cross-sectional study (Gladstone & Koenig, 1994) that has explored different mechanisms to explain different rates of gender differences between a university sample and a high school sample found that high school females reported more depressive symptoms than high school males, whereas university males and females did not differ on depressive symptoms. By examining depression histories over time, it is possible to ascertain whether the inconsistent results in studies of gender differences in depression among university students are due to environmental factors during the university years that may equalize rates of depression among males and females or, alternatively, to selection factors that may favor the admittance of females who are relatively not depression prone into universities or the admittance of high-depression-prone males (see also Nolen-Hoeksema, 1990). Consequently, the study’s second goal is to utilize a prospective longitudinal approach to examine developmentally both the prior depression histories before entering a university (i.e., before age 18) as well as depression rates during the university years (i.e., ages 18–21) for the student and nonstudent groups.

Explanations for the Gender Difference in Depression

Although a comprehensive review of the hypothesized explanations for the occurrence of gender differences in depression is beyond the scope of this article (see Nolen-Hoeksema, 1990; Nolen-Hoeksema & Girdus, 1994, for reviews), we will examine the possibility that the gender difference in depression results from a greater recurrence of depression among females compared with males. The recurrence of depression hypothesis suggests that equal proportions of males and females develop first episodes of depression, but females who have been depressed at least once are more likely than once-depressed males to have recurrences. Thus, it is the greater recurrence of depression among once-depressed females compared with males that would explain the gender difference in depression at any point in time.

Research investigating recurrence after recovery from depression in general has produced mixed and inconsistent results regarding gender. Two studies have found support for gender differences in depression recurrence. The first study (Amenson & Lewinsohn, 1981), which utilized structured clinical interviews in an adult sample, reported that women with a history of depression were more likely than men with a history of...

---

1 Throughout the article we use the term “overall rate of depression” to refer to the 1-year point prevalence rate of clinical depression at the age assessed. We use this term because the overall rate of depression includes both new cases of depression and possible recurrences, and we want to distinguish clearly new cases of depression that develop at a particular age from the 1-year prevalence rate at that age.
In conclusion, this study has three main goals. First, we will describe the development of clinical depression in males and females over a 10-year span from preadolescence (age 10-11) to early adulthood (age 20-21) and pinpoint when the gender difference in depression emerges for both overall rates of depression and new cases of depression. Second, we will construct this developmental timeline separately for a student group and a nonstudent group to investigate whether the gender difference in rates of clinical depression in a student group are comparable to those in a nonstudent group from preadolescence to early adulthood. Third, we will examine whether the gender difference in depression may be explained, at least in part, by greater recurrence of depression or more severe depressive episodes for females. To our knowledge, this study is the first to investigate the cause, onset, and recurrence of clinical depression in a prospective fashion before adulthood (Keller, 1988; Weissman, 1990). Finally, and perhaps most important, by delineating the emergence of gender differences in depression we also construct a general developmental timeline of the course of clinical depression from preadolescence to early adulthood that highlights exactly when the peak increase in rates of depression occurs and how the course of depression evolves over this 10-year period in the early life course.

**Method**

**Participants**

Participants in this study consisted of members of a complete birth cohort from the Dunedin Multidisciplinary Health and Development Study. The history of the sample has been described in detail by Silva (1990). Briefly, this longitudinal study has assessed the health, development, and behavior of members from a complete cohort born between April 1, 1972, and March 31, 1973, in Dunedin, New Zealand. At delivery, perinatal data were obtained. Three years later when these children were followed, 1,037 children (52% male; 91% of eligible births) were included in the longitudinal study and assessed. This formed the base sample. Participants’ fathers are representative of the social class and ethnic distribution of the general population of New Zealand’s South Island. Participants are predominantly of European ancestry, and fewer than 7% of the sample identified themselves at age 18 as Maori or Polynesian. These demographics match the ethnic identity of the South Island of New Zealand. Recent cross-national comparisons (Costello, 1989; Kessler et al., 1994; Newman et al., 1996) show that the prevalence rates for psychiatric disorders, and depression specifically, between the United States and the Dunedin sample are comparable. On the basis of these cross-national comparisons, we believe that the findings relating to mental health from the Dunedin sample can be generalized to the United States and other industrial countries.

The current study used the diagnoses of depression data from ages 11, 13, 15, 18, and 21. Reliabilities of diagnoses (kappa coefficients averaged over disorders) at the different ages were all above .80. In the Age 11 follow-up, 792 (86%; 416 male) of the original study members were interviewed (see Anderson et al., 1987, for details). For the Age 13 follow-up, 728 study members (86%; 378 male) were interviewed (see Frost, Moffitt, & McGee, 1989, for details). At the Age 15 follow-up, 943 (95%; 482 male) were interviewed (see McGee et al., 1990, for details). For the Age 18 assessment, 930 study members (91%; 475 male) were interviewed (see Fehan, McGee, Nada Raja, & Williams, 1994, for details). In the most recent Age 21 follow-up, 961 (97%; 490 male) members were interviewed (see Newman et al., 1996, for details). As can be seen, different numbers of participants were available for interview at various ages. Because this study aims to delineate the development of gender differences in depression over these 10 years, we only include the 653 members (63% of total sample; 334 male) of the data set for whom complete diagnosis of depression data were available for all five assessment ages.

To investigate whether the sample of 653 study members that we included was significantly different from the 384 participants not included, we compared the two samples on rates of depression, depressive symptoms, social class, and IQ by t test. First, there was no differential attrition for males versus females on any of these variables. Next, the two groups were not different on the number of depressive symptoms reported for ages 18 and 21, on rates of depression from ages 18 and
21 (ages 18 and 21 were used because these two ages had the most number of people interviewed), and on social class (all is less than 1). However, the two groups were different on the Wechsler Intelligence Scales for Children at age 7, at (948) = 4.08, p < .001 (108.2 vs. 104.1), and age 11, at (915) = 2.34, p < .05 (108.9 vs. 106.3). In both cases, the study members not included in the final sample had slightly lower mean IQ scores. Finally, we conducted a chi-square analysis to see if those participants not included were more likely to be depressed at any point in their life than those included. This analysis was not significant, \( \chi^2 (1, N = 653) = 0.12 \), indicating that there were equal proportions of lifetime prevalence of depression for the included and not-included samples.

Data Collection Procedure

Over the years, the Dunedin birth cohort has been assessed with a variety of psychological, medical, and sociological measures at ages 3, 5, 7, 9, 11, 13, 15, 18, and 21. The basic procedure includes bringing each study participant into the research unit within 60 days of his or her birthday for a complete day of individual data collection. Interview data have been supplemented by searches of official records at each assessment wave, and questionnaires have been mailed to informants who know the study participants well. In each assessment, informed consent was obtained, and study members were free to decline participation in any portion of the assessment.

Measures

Assessment of depression at ages 11, 13, and 15. The mental health interviews were obtained in private, psychologist- or psychiatrist-conducted interviews. Diagnoses of major depressive episode (MDE) and dysthymia (DD) at age 11 were determined by using the Diagnostic Interview Schedule for Children, Child Version, Version XIII-III (DISC-C). This instrument was developed for the National Institute of Mental Health (Costello, Eydelbrock, Kahns, Kessler, & Klaric, 1982). The DISC-C is a very structured interview based on the Diagnostic and Statistical Manual of Mental Disorders (3rd ed.; DSM-III; American Psychiatric Association, 1980) criteria. Study participants answered these structured questions using a reporting period for the past 12 months. Because a 1-year reporting period was used, we note here that participants who were interviewed at age 11 were reporting on depression from age 10 to 11, at age 13 they were reporting on depression from age 12 to 13, and so on. Throughout this article we will use “age 11” to refer to the participant’s 11th year. Verbal responses from the study members were recorded as no (scored 0), sometimes (scored 1), and yes (scored 2). Only those responses of yes were deemed severe enough to be entered into the diagnostic algorithm. Depression diagnoses were determined by algorithms following DSM-III criteria. Impairment criteria were not used as part of the diagnostic algorithms because DSM-III did not require impairment for a diagnosis. Reliability and validity of the DISC-C in this sample for ages 10–11 are described elsewhere (see Anderson et al., 1987). The DISC-C was also used at age 13 and 15, although minor modifications were made from the version used at ages 10–11. Details of the modifications and the psychometric properties can be found elsewhere (see Frost et al., 1989, for age 13; and McGee et al., 1990, for age 15). The important modifications relevant to this study included using a “gating” procedure for depressive disorders such that depressed mood or anhedonia and time criteria for the duration of depressive symptoms had to be met as initial criteria before any further depression questions (i.e., depressive symptoms) were asked. We believe it is highly unlikely that this gating procedure dramatically changed the rate of diagnosed clinical depression for two reasons. First, as stated previously, the prevalence rates of psychiatric disorders in the United States and New Zealand are very comparable (Costello, 1989). Second, the initial gating criteria of depressed mood and duration are required to make a diagnosis for MDE and DD. In summary, DSM-III criteria were used to arrive at a diagnosis of clinical depression (MDE and DD combined) for ages 11, 13, and 15 using the DISC-C with a reporting period of 1 year.

Assessment of depression at ages 18 and 21. Diagnoses of MDE and DD at ages 18 and 21 were determined by using a modified version of the Diagnostic Interview Schedule, version III-R (DIS: Robins, Helzer, Croughan, & Ratcliff, 1981). These modifications and the psychometric properties for each age assessment are described in detail elsewhere (see Feehan et al., 1994, for age 18; and Newman et al., 1996, for age 21). Briefly, the alterations included the following: (a) using DSM-III-R (3rd ed., rev.; American Psychiatric Association, 1987) criteria only; (b) using a reporting period for the past 12 months; (c) limiting the response options to 0, 1, and 2 as described above; and (d) using a gating procedure requiring depressed mood or anhedonia and duration of depressive symptoms as described above. Thus, study participants were given a diagnosis of clinical depression (MDE and DD combined) for the age 18 and 21 assessments using the past year as a reporting period if they met requisite DSM-III-R criteria. As with the DISC-C, impairment criteria were not used as part of the diagnostic algorithms to form depression diagnoses because DSM-III-R did not require impairment as part of the diagnosis.

Because the principal goal of this study is to document the development of depression over the span of 10 years from preadolescence to young adulthood, we note here that the depression items for the DISC-C and DIS are very similar. The only changes made from the DISC-C to the DIS include changing the wording in the DIS to make it more appropriate for adults. We highlight that these two instruments are nearly identical to illustrate that any differences seen using the DISC-C and the DIS are likely to be due to age (development) effects and not to effects of instrument.

Assessment of depressive symptoms. The scales for depressive symptoms were computed by summing answers to the various depression symptom questions from the DISC-C and DIS. Because of the gating procedure used for ages 13, 15, 18, and 21 (described above), participants who did not pass the gate would receive a score of zero for the symptom scale. The depressive symptoms questions at ages 13, 15, 18, and 21 were asked only if the respondents met the initial criteria for depressive mood and duration. This methodology likely results in an accurate representation of the number of depressive symptoms for participants at age 11 but an underrepresentation of the potential number of symptoms at ages 13, 15, 18, and 21 among participants who did not meet the gating criteria. Finally, some participants were missing depressive symptom scales for the Age 11 (n = 35) assessment. The primary reason for missing scale scores at this age is that scales were not computed and archived for those study participants who adamantly answered “I don’t know” to any specific depressive symptom questions.

Assessment of main occupation at age 21. The second goal of the study is to construct a developmental timeline for depression separately for a sample of university students and nonstudents. Therefore, information collected on the participants’ main activity at age 21 was used toward this goal. For the sample of 653 participants, 13% were unemployed (61% male), 53% were employed (55% male), 0.5% were in a government training program (0% male), 29% were students currently enrolled in a university or 4-year college (48% male), and 5% were homemakers (3% male). The participants who were unemployed, employed, in a training program, and homemakers comprised the nonstudent group, whereas currently enrolled students formed the student group.

The degrees of freedom for the two comparisons at ages 7 and 11 are different because there were different numbers of participants who were administered the Wechsler Intelligence Scales for Children at those two ages.
Temporal Resolution of Gender Differences in Depression

Development of clinical depression. The first set of analyses was directed at pinpointing when the gender difference in clinical depression (MDE and DD combined) emerges and describing how depression develops separately by gender for the 10-year span from age 11 to age 21. Table 1 shows the descriptive statistics for the diagnosis of depression data for all ages by gender. Figure 1 depicts the development of rates of diagnosed depression (1-year point prevalence) at each assessment period for the whole sample and for gender separately over the 10-year period. Visual inspection of this figure reveals that the rates of depression appeared to increase over time for both males and females. Further, small gender differences in depression first seemed to emerge after age 13 when females' rates continued to rise, whereas males' rates stayed steady. Finally, the rates of depression rose rapidly after age 15 for both males and females, with females demonstrating an even steeper rise than the males. Thus, although first apparent after age 13, gender differences in depression did not become dramatic until after age 15.

To verify this visual interpretation, a repeated measures analysis of variance (ANOVA) was performed on the diagnosis of depression data with Gender as the between-subjects factor and Age as the within-subjects factor. The ANOVA produced significant main effects for Gender, \( F(1, 651) = 29.94, p < .001 \); and Age, \( F(4, 2604) = 75.96, p < .001 \); and a significant Gender \( \times \) Age interaction, \( F(4, 2604) = 13.23, p < .001 \). We also redid these analyses using Cochran's Q statistic (Winer, 1971, pp. 303–305) for repeated measures with dichotomous data in case the dichotomous nature of the depression diagnosis dependent variable seriously violated assumptions of the ANOVA. Consistent with the results from the ANOVA above, all results were highly significant: Gender, \( \chi^2(1, N = 653) = 41.6, p < .001 \); Age, \( \chi^2(4, N = 653) = 304.75, p < .001 \); and Gender \( \times \) Age interaction, \( \chi^2(4, N = 653) = 53.14, p < .001 \).

Following Keppel's (1991, pp. 383–388) suggestion, we performed tests on simple main effects of Gender and Age to decompose the significant Gender \( \times \) Age interaction. Simple main effects tests for Age showed that the rates of depression increased over time for females, \( F(4, 1272) = 56.67, p < .001 \); and for males, \( F(4, 1332) = 19.37, p < .001 \). These simple main effects tests were analyzed further by simple comparisons (paired \( t \) tests) that compared each of the continuous age periods. These results show that the females' rates of depression rose from ages 11 to 13 and from ages 15 to 18 and were steady from ages 13 to 15 and 18 to 21. The males' rates rose only between ages 15 to 18, whereas they remained relatively constant from ages 11 to 15 and from ages 18 to 21. Finally, depression rates rose significantly for the whole sample after age 15 but then became level after age 18.

The simple main effects tests for Gender involved a series of

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lifetime Depression Occurrence by Gender</strong></td>
</tr>
<tr>
<td>Factor</td>
</tr>
<tr>
<td>One-year point prevalence of overall depressive cases</td>
</tr>
<tr>
<td>% depressed at age 11</td>
</tr>
<tr>
<td>% depressed at age 13</td>
</tr>
<tr>
<td>% depressed at age 15</td>
</tr>
<tr>
<td>% depressed at age 18</td>
</tr>
<tr>
<td>% depressed at age 21</td>
</tr>
<tr>
<td>Incidence of new depressive cases</td>
</tr>
<tr>
<td>% first depressed at age 11</td>
</tr>
<tr>
<td>% first depressed at age 13</td>
</tr>
<tr>
<td>% first depressed at age 15</td>
</tr>
<tr>
<td>% first depressed at age 18</td>
</tr>
<tr>
<td>% first depressed at age 21</td>
</tr>
<tr>
<td>Lifetime prevalence of depression (at least one episode of depression)</td>
</tr>
<tr>
<td>% up to age 11</td>
</tr>
<tr>
<td>% up to age 13</td>
</tr>
<tr>
<td>% up to age 15</td>
</tr>
<tr>
<td>% up to age 18</td>
</tr>
<tr>
<td>% up to age 21</td>
</tr>
<tr>
<td>Number of depressive episodes in life</td>
</tr>
<tr>
<td>% with 0 episodes</td>
</tr>
<tr>
<td>% with 1 episode</td>
</tr>
<tr>
<td>% with 2 episodes</td>
</tr>
<tr>
<td>% with 3 episodes</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
DEVELOPMENT OF DEPRESSION

Figure 1. Development of overall rates of clinical depression (1-year point prevalence combining new cases and recurrences) by age and gender.

independent t tests with Gender as the between-subjects factor at each age. Males showed a tendency to have a higher rate of clinical depression at age 11 than females, females, $t(651) = 1.84, p < .06$ (males 2% vs. females 0%); males and females did not differ at age 13, $t(651) = .08, ns$ (2% each); females had higher rates at age 15, $t(651) = 2.49, p < .01$ (1% vs. 4%); at age 18, $t(651) = 2.49, p < .001$ (11% vs. 23%); and at age 21, $t(651) = 4.73, p < .001$ (11% vs. 25%). These results demonstrate that the gender difference in clinical depression in this sample began to emerge between ages 13 and 15. The males tended to have higher rates of depression at age 11, males and females had virtually identical rates at age 13, and then the females had higher rates from age 15 to age 21.

In addition to these simple effects, we also conducted several smaller ANOVAs with Gender as the between-subjects factor and two continuous ages as the within-subjects factor as another way of breaking down the Age × Gender interaction and of pinpointing the time period when the gender difference in depression emerges. The analyses for the Age × Gender interaction for the two continuous ages were nonsignificant for ages 11 to 13 and for ages 18 to 21. A tendency toward significance was found for ages 13 to 15, $F(1, 651) = 3.65, p = .056$; and a significant result was found for the period between ages 15 and 18, $F(1, 651) = 9.46, p < .01$. These results verify the visual interpretation of Figure 1: The gender difference in depression began to emerge somewhat after age 13 but became very noticeable and significant after age 15 with more females exhibiting depression than males.

Student versus nonstudent groups. The second set of analyses examined whether the gender differences in depression are present in both the university student group and the nonstudent group. Table 2 shows the descriptive statistics for both the student and nonstudent group. Figure 2 shows the development of clinical depression (MDE and DD combined) at each of the ages by gender for the student and nonstudent group separately. To determine whether student status had any impact on the development of gender differences in clinical depression, we performed a repeated measures ANOVA on the diagnosis of depression data with Gender and Student Status as the between-subjects factors and Age as the within-subjects factor. Significant main effects were found for Gender, $F(1, 649) = 27.36, p < .001$; and Age, $F(4, 2596) = 62.86, p < .001$; and the Gender × Age interaction was significant, $F(4, 2596) = 13.27, p < .001$. The main effect for Student Status was nonsignificant, $F(1, 649) = .11$; and Age × Student Status, $F(4, 2596) = .12$; and Age × Gender × Student Status, $F(4, 2596) = 1.06$. We also redid these analyses using Cochran's Q statistic and found the same pattern of significant and nonsignificant results. Simple main effects were not computed because Student Status did not affect the diagnosis of depression data, and the appropriate significant interactions were already decomposed in the previous set of analyses. Thus, these results show that the pattern of gender differences in rates of clinical depression was similar (i.e., not statistically different) for the student and nonstudent groups.

Recurrence of Depression Hypothesis

This series of analyses was conducted to ascertain whether the gender difference in depression obtained in the above analyses could be explained by females having more depression recurrences than males. Because various components comprise the overall pattern of gender differences in depression (e.g., new cases and recurrence), we separated these factors and analyzed each independently to see whether each alone or in combination might explain the gender difference in depression. First, we looked at the development of new cases of depression over the various ages. Second, we tested for gender differences in depression recurrence.
Table 2

**Lifetime Depression Occurrence by Gender and Student Status**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Student</th>
<th></th>
<th></th>
<th></th>
<th>Nonstudent</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Total</td>
<td>Male</td>
<td>Female</td>
<td>Total</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>One-year point prevalence of overall depressive cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% depressed at age 11</td>
<td>1.1</td>
<td>0.0</td>
<td>0.53</td>
<td>2.0</td>
<td>0.4</td>
<td>1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% depressed at age 13</td>
<td>1.1</td>
<td>0.0</td>
<td>0.53</td>
<td>2.4</td>
<td>3.1</td>
<td>2.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% depressed at age 15</td>
<td>1.1</td>
<td>2.0</td>
<td>1.59</td>
<td>1.2</td>
<td>5.4</td>
<td>3.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% depressed at age 18</td>
<td>8.9</td>
<td>23.5</td>
<td>16.49**</td>
<td>11.4</td>
<td>23.3</td>
<td>16.99***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% depressed at age 21</td>
<td>7.8</td>
<td>27.5</td>
<td>18.08***</td>
<td>12.3</td>
<td>23.9</td>
<td>17.85***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .06.  **p < .01.  ***p < .001.

**Development of new cases of depression.** The pattern of new cases of depression is important to the recurrence of depression hypothesis because the recurrence hypothesis states that equal proportions of males and females will experience a first (new case) depressive episode. Here, we define new case incidence as the age at which the individual is first diagnosed with an episode of depression. Thus, someone diagnosed with depression at age 11 and again at age 15 would be counted only once at age 11. Table 1 shows the descriptive statistics for prevalence of new depression cases. Figure 3 shows the relation between new incidence of depression, age, and gender.

To examine the new incidence of depression, we conducted another repeated measures ANOVA with Gender as the between-subjects factor and Age as the within-subjects factor on the dependent variable of new cases of depression at each age. Individuals who never received any clinical diagnoses of depression over the course of the study were coded as “0” at all ages. Individuals who received a clinical diagnosis of depression were coded as “1” for the age at which they first received a depression diagnosis, whereas they were coded as “0” at all other ages regardless of whether they developed an additional depression diagnosis at a later age. Results demonstrated a main effect for Gender, *F*(1, 651) = 38.63, *p* < .001; for Age, *F*(4, 2604) = 44.08, *p* < .001; and a Gender × Age interaction, *F*(4, 2604) = 8.03, *p* < .001. Reanalysis using the nonparametric Cochran’s Q statistic yielded similar significant results. Finally, we also entered Student Status and Gender as between-subjects factors with Age as the within-subjects factor to see whether nonstudents have a similar rate of new depression cases compared with students. This repeated measures ANOVA yielded nonsignificant results for Student Status, *F*(1, 649) = .65; for Gender × Student Status interaction, *F*(1, 649) = .26; for Age × Student Status interaction, *F*(4, 2596) = .31; and for the Age × Gender × Student Status interaction, *F*(4, 2596) = .95; and significant results for Gender, *F*(1, 649) = 35.21, *p* < .001; Age, *F*(4, 2596) = 38.34, *p* < .001; and Age × Gender interaction, *F*(4, 2596) = 7.99, *p* < .001.

Simple main effects for Age showed that the rates of new cases...
Development of Depression

The development of new cases of clinical depression (1-year point prevalence) by age and gender.

Figure 3. Development of new cases of clinical depression (1-year point prevalence) by age and gender.

The rates of new incidence of depression rose significantly from ages 11 to 13 and from ages 15 to 18, but they were otherwise level from ages 13 to 15 and from ages 18 to 21. The males' new cases of depression remained steady from each age to the next except for an increase between ages 15 to 18. Finally, the whole sample showed an increase in new cases of depression between ages 15 to 18 and a decrease from ages 18 to 21, whereas these rates were even from ages 11 to 15.

The simple main effects for Gender included independent t-tests with Gender as the between-subjects factor at each age with new case incidence of depression as the dependent variable. Females demonstrated significantly greater increases in new cases of depression compared to males at age 15, t(651) = 3.15, p < .01; age 18, t(651) = 4.02, p < .001; and age 21, t(651) = 3.52, p < .001. These results show that significantly more females than males were becoming depressed for the first time from age 15 to age 21.

Again, we conducted several smaller ANOVAs with Gender as the between-subjects factor and two continuous ages as the within-subjects factor to break down the Age X Gender interaction and to understand better when the gender difference in new depression cases emerges. Results were nonsignificant for ages 11 to 13 and for ages 18 to 21, whereas significant results were found for ages 13 to 15, F(1, 651) = 4.86, p < .05; and for ages 15 to 18, F(1, 651) = 5.53, p < .05. These analyses show that more females compared to males were becoming depressed for the first time between the ages of 13 to 18.

Greater recurrence of depression. The recurrence hypothesis states that equal proportions of males and females will experience at least one episode of depression (the first episode), but that more females than males will go on to have greater than one lifetime episode. Among those participants having at least one depressive episode over the course of the study, the chi-square analysis examining the relationship between gender and having only one lifetime episode versus greater than one lifetime episode was not significant, χ²(1, N = 205) = 1.95 (see Table 1 for overall depression recurrence data). Specifically, of those participants with at least one lifetime depressive episode, 74% of males had only one lifetime depressive episode compared to 74% of females, whereas 26% of males had greater than one depressive episode compared to 26% of females. Thus, this analysis does not support the hypothesis that once depressed, females are more likely than males to experience recurrences of depression in this period of the early life course. We caution the reader that these recurrence rates should not be interpreted as population recurrence rates because we have only followed the participants up to their 21st year and have assessed 1-year point prevalence rates every other year so that some participants may have experienced a depressive episode in a year that was not recorded.

Symptom Severity Analyses

The following series of analyses involves using the depressive symptom scales from the DISC-C and DIS reported by study participants who received a clinical diagnosis of depression as a measure of the severity of those participants' depressive episode or episodes (where each symptom was scored 0, 1, or 2). In conducting these severity analyses, we only considered the depressive symptoms reported by those individuals who were diagnosed as depressed.

For the first set of analyses looking at average symptom severity per episode, we used one symptom severity score for those participants who had at least one episode of depression. Thus, for those participants with more than one depressive episode, we computed an average symptom severity score from the multiple
episodes and used that average symptom score as the measure of symptom severity. Results of this analysis were nonsignificant, \( t(203) = 0.594 \), with males reporting a mean symptom score of 21.4 (SD = 9.06) symptoms versus females' mean of 22.1 (SD = 7.66) symptoms.

The second set of analyses looked at the symptom severity for new cases of depression. For these analyses, someone who was diagnosed as depressed at ages 13 and 18 would only be counted once for their first episode (new incidence) at age 13. A series of independent \( t \) tests was computed with Gender as a between-subjects variable on the dependent measure of depressive symptoms reported for individuals' first depressive episode. Out of all the ages, only age 15 approached significance, \( t(14) = 2.06, p = .08 \), with females reporting a mean score of 24.42 (SD = 10.48) compared with 8.0 (SD = 11.31) for males. Analyses for the other ages were nonsignificant. Finally, pooling the symptom scores together for all ages from the first episode also failed to show significant differences by gender, \( t(203) = 0.45 \), with males reporting a mean score of 22.07 (SD = 10.57) for their first depressive episode compared to females' 21.49 (SD = 7.68).

Finally, we examined gender differences for the most severe episode of depression. For participants with only one lifetime depressive episode, depressive symptom scores from the one episode were used, whereas the highest depressive symptom score from the specific age was chosen for participants with multiple lifetime episodes. An independent \( t \) test with Gender as a between-subjects variable on the dependent variable of the number of depressive symptoms reported for the most severe depressive episode was nonsignificant, \( t(203) = 0.21 \), with males' mean symptom score of 23.55 (SD = 11.72) compared with females' 23.22 (SD = 9.74) for most severe episode. Thus, symptom severity analyses for the average depressive episode, for first episodes of depression, and for the most severe depressive episode did not support the hypothesis that females who received a clinical diagnosis of depression reported more depressive symptoms than males with a clinical diagnosis of depression.

**Discussion**

The central goal of the study was to describe the emergence of gender differences in clinical depression and the overall development of depression from preadolescence to young adulthood. The development of depression and the gender differences can be seen most clearly in Figure 1 for the overall rates and in Figure 3 for the new cases of depression. The beginning of the emergence of the gender difference occurs between ages 13 and 15 for both the overall rates of depression as well as new cases of depression. However, we believe the period from ages 15 to 18 is the most important age range to focus on not only because the rates of depression increase dramatically for both genders, but also because the female rate of depression rises to double the prevalence rate for males. This dramatic increase is most striking, perhaps, because the overall rates do not continue to increase from age 18 to 21, and the new cases of depression actually begin to decline. Thus, mid-to-late adolescence may represent a critical time period for heightened vulnerability to depression, especially new incidence of depression that is greater than the risk associated with childhood, and possibly even adulthood. This heightened vulnerability from ages 15 to 18 is especially great among females.

In comparison with previous studies, our finding that gender differences in clinical depression begin to emerge from ages 13 to 15 is in agreement with Ge et al.'s (1994) and Petersen et al.'s (1991) finding the emergence of a gender difference in depressive symptoms around age 13. Our results regarding the jump in depression rates from age 15 to 18 are similar to findings from the retrospective reports gathered in the Epidemiological Catchment Area study (Burke, Burke, Regier, & Rae, 1990). However, our finding of a large increase in depression rates for the whole sample and particularly for females from age 15 to 18 is different from Lewinsohn et al.'s (1993) results of no overall increase in depression from age 14 to 18 and no Age \( \times \) Gender interaction.

It is difficult to determine what factors might explain the discrepancy in results. The use of different diagnostic interviews may have produced different age and gender results. However, we believe this is unlikely because both investigations yielded very similar lifetime prevalence rates for the same age groups (up to age 18), with Lewinsohn et al.'s (1993) finding that there is a 27% lifetime prevalence rate for females and a 13% rate for males, whereas our results showed a 28% lifetime prevalence rate for females and 14% for males. A more likely explanation for the discrepancy in results may be the difference in methodology. Lewinsohn et al. (1993) used a design in which diagnostic interviews were given to the participants of all ages 14–18 at Time 1 and then again a year later at Time 2. On the other hand, we used a prospective longitudinal design in which the participants were interviewed five times over the course of 10 years with a reporting period of the past year. Lewinsohn et al.'s (1993) design of a 1-year follow-up period may not have been enough time for the subtle age and gender effects in depression to occur. Ultimately, the ideal methodology to obtain accurate depression incidence data is a prospective longitudinal design using large samples of participants (Eaton et al., 1989; Lewinsohn et al., 1993) because the incidence of depression is low, especially at earlier ages. Thus, we believe that our prospective longitudinal study covering 10 years with a large sample of participants is a particularly powerful design that enables the detection of subtle age and gender differences in depression.

The second goal of the study was to determine whether the pattern of gender differences in clinical depression found in the overall sample would also be obtained when the sample was broken down into a group of students attending university and a group of nonstudents. Results showed that the development of clinical depression and new cases of depression over the 10-year period from preadolescence to young adulthood is very similar for males and females in both university student and nonstudent groups drawn from the same birth cohort. Students attending universities in New Zealand are quite similar to those attending universities in the United States. These results are important not only because they shed light on the true pattern of gender differences, but also because they provide compelling evidence that the results from many depression studies using university students may generalize to the general population at the same age. We also hoped to elucidate plausible reasons why previous studies examining gender differences in depressive
symptoms in student groups have yielded inconsistent results. The results of no differences in the rates of depression for students compared to nonstudents from age 11 to age 21 rule out the possibility that selection factors were operating before entry into university that would have favored the admittance into universities of either less depressed females or more depressed males.

To resolve this inconsistency, we suggest that our finding of a clear gender difference among university students, compared with some previous studies that have not found this result, may be explained by the difference in methodologies. Our diagnostic interviews assessed participants' depression for the past year, whereas the other studies administered a depressive symptoms questionnaire assessing depression for a much shorter time frame (typically 1 day to 1 week). We would expect more inconsistent results from studies that are using a shorter time frame to assess depression because these studies were, in essence, taking a "snap shot" of depression for a brief time period. In contrast, at each assessment, our method "aggregates" (Epstein, 1980) depression over 365 days and, thus, is more likely to reveal the true pattern of gender differences in depression.

The third goal of the study was to examine potential explanations of the gender difference found in clinical depression, such as the recurrence hypothesis and females experiencing more severe depressive episodes. For the recurrence hypothesis to explain completely the gender difference in depression, it is necessary to show that (a) equal proportions of males and females experience first time depressive episodes (new cases); and (b) females experience more recurrent depressive episodes than males. Contrary to the recurrence hypothesis, we found that twice as many females as males experienced a new case of depression and that there was no gender difference in depression recurrence during the 10 years of the study among participants who had developed at least one depressive episode. Thus, no support was found for the hypothesis that the gender difference in depression could be explained by more depression recurrences among females compared with males up to the assessment at age 21.

Of interest, we found almost identical results for overall depression rates and new cases of depression. This finding does not suggest that there were no depression recurrences because, among participants with at least one lifetime depressive episode, 25% had experienced a recurrence. It is important to keep in mind that our participants are still young, and that given more time, about 50% of participants who experienced a first depressive episode are likely to reexperience another episode (Belsher & Cosciello, 1989). It is possible that greater female recurrence of depression will occur at later ages, although data from the National Comorbidity Study suggest this is unlikely (Kessler et al., 1994). We suggest that further research be conducted using older samples to examine more closely the inconsistent adult results regarding gender differences in depression recurrence. Regarding gender differences in depressive symptom severity, results showed that males and females who had been diagnosed as ever having a depressive episode did not differ in the average number of symptoms they reported across episodes, in the number of symptoms during their first depressive episode, or in the number of symptoms for their most severe episode.

Research on depression recurrence from childhood to adulthood, especially using longitudinal studies, is rare (Gotlib & Hammen, 1992). Furthermore, the majority of research on depression recurrence has followed depressed clinic inpatients over time to ascertain whether they experience a recurrence. Yet, the strategy of using only clinic samples to understand depression recurrence may be problematic and suggestive of greater continuity in depression than really exists because individuals who seek help in clinics tend to have a longer duration of illness and have a more negative prognosis (Cohen & Cohen, 1984). Thus, the evidence from this investigation is important in establishing that about 25% of individuals will likely experience another depressive episode from midadolescence to young adulthood when prospectively assessed in an epidemiological sample. These results support previous longitudinal research in clinics (Garber, Kriss, Koch, & Lindholm, 1988; Harrington, Fudge, Rutter, Pickles, & Hill, 1990), in which individuals who were treated for depression during adolescence were found to be at increased risk of experiencing a recurrence when contacted later in adulthood. However, not everyone with a previous depressive episode goes on to reexperience another depression. Future research should examine factors that contribute to both the vulnerability and invulnerability in depression recurrence over time with a special focus on the heightened risk to depression that exists during middle-to-late adolescence.

Directions for Future Research

Regarding the large increase in depression rates between ages 15 and 18, vulnerability-stress models of depression offer a framework to understand how cognitive vulnerability (Abramson, Metalsky, & Alloy, 1989) and interpersonal vulnerability (Gotlib & Hammen, 1992) may interact with the increasing levels of stress experienced by youth progressing through adolescence. Results from a longitudinal investigation (Ge et al., 1994) highlight the importance of stressful events in risk for depression, especially between ages 13 to 18 when the number of reported stressful life events increases dramatically. Second, understanding whether the pattern of development of depression over time is unique or overlaps with that of other psychiatric disorders is especially important given the known comorbidity of depression with other psychopathology, especially anxiety (Brady & Kendall, 1992; Kovacs, Gatsonis, Paulauskas, & Richards, 1989) and conduct disorders (Kovacs, Paulauskas, Gatsonis, & Richards, 1988). Researchers have found comorbidity of disorders is associated with increased risk for continued depression and recurrence (Brady & Kendall, 1992). Also, females tend to be more anxious-depressed, whereas males tend to be more conduct disordered-depressed (Craighead, 1991), suggesting that there are gender differences in comorbidity with depression. Third, researchers need to be sensitive to age differences when choosing age groups to examine developmental differences in depression because of the large increase in depression rates from age 15 to 18. The age and gender differences suggest that different causal processes may be operating at different ages for males and females, or alternatively, that the same causal factors may be operative across time and gender but increase in intensity over time, especially for females, to cause the greater rates of depression in later adolescence. Finally, given
the dramatic increase in depression rates and the high risk of depressive recurrence later in adulthood (Harrington et al., 1990), therapeutic and intervention resources should target individuals in the 14- to 18-year age group and more research should investigate potential treatment programs specific to these ages (see Birmaher, Ryan, Williamson, Brent, & Kaufman, 1996, for a review of depression treatment in youth).

Limitations

Some cautions regarding this study should be considered. The first concerns the generalizability of results with participants from New Zealand to the United States. Although there are certainly some cultural differences between the two countries, it is doubtful that using a New Zealand sample is problematic for research on development of depression because cross-national comparisons (Costello, 1989; Kessler et al., 1994; Newman et al., 1996; Weissman et al., 1996) have shown that the rates of psychiatric diagnoses are very similar between the two countries. Further, the rates of depression for this sample are very similar to those reported elsewhere for similar age groups (Lewinsohn et al., 1993). Regarding gender differences in depression, cross-national studies have shown that the gender difference in depression is a robust finding across many countries (Nolen-Hoeksema, 1990). In fact, using New Zealand participants may be an advantage over some studies conducted in the United States that tend to rely heavily on participants from research universities. Thus, results from our study may be quite generalizable because we used a large sample of representative, unselected members from the general population.

The second caution involves the use of children and adolescents reporting on their own depression by use of DISC-C and DIS. Research has shown that children can accurately report on their depressive affect and depressive symptoms (Kazdin, 1994) and that children after age 10 can readily recognize emotions (Kovacs, 1986). Nevertheless, using multiple informants, such as teachers and parents, in addition to child’s report, is the ideal method, and previous studies using this Dunedin data set have validated the potentially more suspect diagnoses at earlier ages by parent and teacher reports (Anderson et al., 1987). In addition, interviews are the best assessment method and provide the most accurate data when children are honest and trust the interviewer (Kovacs, 1986). We believe that the Dunedin data are ideal in this sense because the study members have been assessed multiple times in the past and have provided candid reports on sensitive issues, without violations of confidentiality. This combination provides additional confidence that these are accurate reports of depression diagnoses. Next, in using both the DISC-C and DIS, it is possible that age-related changes observed over time are due to the switch in instruments and not to real developmental effects. However, this possibility is unlikely because other longitudinal studies (Ge et al., 1994; Peterson et al., 1991) found an increase in depressive symptoms over adolescence that closely parallels our findings for diagnoses of clinical depression. Finally, using self-report methodology by DISC-C and DIS interviews raises the possibility that the gender difference in depression may be due to reporting tendencies in which females are more likely than males to reveal and discuss depressive symptoms. However, this reporting bias does not seem likely because prior studies have shown that males of different ages are equally as likely as females to report and discuss emotional problems such as depression (Amenson & Lewinsohn, 1981; Nolen-Hoeksema, 1990) and because there are gender differences (males greater prevalence than females) in other disorders, such as disruptive behavior disorders (Lewinsohn et al., 1993).

The third caution involves how the DISC-C and DIS have been used as interview tools throughout the 10 years of this study. First, participants had to pass the depression ‘‘gate,’’ which inquired about loss of interest and sad mood for a minimum of 2 weeks, before the rest of the depressive symptom questions were asked for ages 13, 15, 18, and 21. We do not view this as a severe limitation because the study primarily focused on depression diagnoses, and a minimum duration of 2 weeks of either sad mood or loss of interest is required by DSM-III and DSM-III-R to diagnose depression. Second, 1-year prevalence estimates were excluded between each age period assessed (e.g., between the age 11 and 13 assessments there are no data on age 11 to 12) so that we do not know whether participants had episodes of depression during these unrecorded years. The results of the study should be considered with these missing years of prevalence estimates in mind.

Conclusion

In conclusion, this investigation used a prospective longitudinal approach with structured diagnostic interviews administered five times over the course of the 10-year study period to examine the emergence of gender differences in clinical depression and the overall development of depression from preadolescence (age 11) to young adulthood (age 21). The gender difference began to emerge between ages 13 and 15, but the most dramatic gender divergence occurred between the ages of 15 and 18. No differences were found on the rates of depression between a university student sample and nonstudent sample. Moreover, the timeline of the emergence of gender differences in depression was highly similar for the two samples. The observed gender difference in depression was not accounted for by females having more recurrences of depression compared to males, and there was no gender difference on depressive symptom severity. Finally, the general developmental timeline describing the course of clinical depression over the 10 years showed that the peak increase in both overall rates of depression and new case incidence of depression occurs between the ages of 15 and 18, suggesting that this age period may represent a critical time in the early life course when risk for depression onset is the greatest, especially among females.

References


Received December 17, 1996
Revision received April 25, 1997
Accepted April 25, 1997