Chapter 2
Problem Behavior of Internationally Adopted Adolescents: A Review and Meta-analysis  

Abstract

In this paper we examine the prevalence of problem behaviors in samples of adolescents who were adopted from a foreign country as infants or young children. We reviewed ten studies and performed a meta-analysis, comparing 2317 internationally adopted adolescents with 14,345 nonadopted adolescents. Results indicate that internationally adopted adolescents exhibit more behavior problems than do nonadopted adolescents ($d = 0.08; p = 0.02$), with the difference seen in externalizing ($d = 0.11; p = 0.00$) but not in internalizing ($d = 0.05; p = 0.12$) behavior problems. Significantly more total behavior problems were seen in adopted than in nonadopted girls ($d = 0.10; p = 0.03$), but not in adopted boys compared to nonadopted ones ($d = 0.07; p = 0.22$). All differences, however, were small. The differences between adopted and nonadopted adolescents were somewhat larger when we considered behavior problems in the clinical range. The majority of the adopted adolescents are well adjusted and do not display significantly more problem behaviors than do their nonadopted peers.

Introduction

International adoption is a relatively new and still-expanding practice. It developed as a consequence of World War II, when many children were orphaned. In the beginning, international adoption involved only a small number of children from relatively few countries, but recent decades have seen a substantial growth in the phenomenon. Today it involves thousands of children and over a hundred countries, whether as states of origin, as receiving

---

states, or both. However, over the decades, the characteristics of the practice have changed substantially.

First, the countries of origin have changed. In the late 1940s, most of the children came from war-torn European nations. These children had lost their parents and family and were sent for adoption to other European countries or the U.S. From 1955 through the early 1980s, most adopted children came from Korea or Vietnam. These children were war orphans or mixed-race offspring of American soldiers. In the 1990s children were adopted mainly from Romania, China, and Russia. The changes in the principal countries of origin have stemmed from crises in these countries (poverty, disease, and famine, as well as war) and shifting attitudes of the government and the population toward international adoption and family planning (birth control, single parenthood).

Second, there have been changes in the number of internationally adopted children. The number of children adopted depends on the situation in the countries of origin, as discussed above, but it is also affected by the situation in the receiving countries, such as the attitude toward international adoption (greater acceptance of international adoption usually increases the rate of adoption) and the possibility of and attitude toward artificial insemination for infertile couples (greater acceptance of artificial insemination may reduce the rate of adoption). In general, however, the number of international adoptions has increased substantially over the last few decades. Today, the number of international adoptions has risen, worldwide, to more than 32,000 a year and will probably increase further.

In addition, parents’ motivation to adopt a child has also changed. In the 1950s and 1960s, the motivation was largely charitable. People adopted because they wanted to provide a home for a child who had been orphaned during the war. Most of these parents already had children of their own. However, since 1970 most applications for international adoption have come from childless couples who cannot have children of their own but still wish to build a family. Most of these adopted children are victims of poverty, rather than of war.

The declining number of people who adopt a child for charitable reasons results in part from the realization that international adoption has a problematic side, too. Many (internationally) adopted children have experienced a number of disadvantageous factors that may influence their adjustment and the parent-child relationship. Some of these factors are pre- or perinatal, whereas others occur after birth. For example, women may suffer stress, malnutrition, or disease during pregnancy and may receive inadequate medical care, any of which can affect the developing fetus. After birth, many children experience (continuing) malnutrition, discontinuous care-taking, poor adult-child relationships, abuse, and lack of both affection and adequate stimulation, as well as poor medical care. These disadvantageous factors are
Problem Behavior of Adopted Adolescents

found especially in internationally adopted children, who come from countries where crises such as war, poverty, disease, and famine prevail, and where negative attitudes toward both international adoption and issues like unwanted pregnancy and single parenthood are common. After the children are adopted, they must become accustomed to a new environment and become familiar with new parents. In the case of international adoption, the new environment (for example the climate) can be very different from the old one. The child and the adoptive parents often have to come to terms with their different appearances, which may complicate the process of reciprocal identification. All of these factors can have long-term negative influences on the lives of both the (internationally) adopted child and the adoptive parents.

Adolescence is probably one of the most difficult periods for adopted children and their parents. Due to a variety of physical and cognitive changes (such as the shift to more-abstract concepts) during this time, adolescents become more preoccupied with thinking about who they are, where they come from, and what they will become. By integrating different thoughts about themselves, they establish a sense of self. Although children learn to understand their racial identity at about the age of 7 and the meaning of adoption at about 8–12 years, they tend to become more concerned with these issues somewhat later. For adopted children, the period of adolescence and especially the process of establishing a sense of self can be very difficult for several reasons.

First, as some adoption theorists have emphasized, establishing a stable sense of self is more complex for adoptees because they have been cut off from their origins and are often prevented from gaining information about their birth heritage. A part of their lives is missing. From clinical practice it is clear that some adopted adolescents feel as if they have lost a part of themselves. Moreover, in thinking about their origins and biological family, adopted adolescents may have to deal with loyalty conflicts toward their adoptive parents, which may make them feel guilty about gaining information about their birth heritage.

Second, adopted adolescents have to cope with the fact that they were given up by their biological parents. Often it is unclear why they were given up, which leaves them feeling confused, tormented, and angry. As a consequence of this lack of clarity, many adopted adolescents create fantasies about the reason for the adoption and about their biological parents’ lives. Some start to idealize their biological parents, often at the expense of their adoptive ones.

Third, it may be difficult for adopted adolescents to identify with their adoptive parents or brothers and sisters. Adoptees often do not know their biological family and (particularly in the case of internationally adopted children) have a different appearance and different traits from their adoptive
family. This makes it difficult for them to come to terms with their identity, and it may make them feel as if they do not really belong to a family.

Finally, in coming to terms with their identity, adolescents often compare themselves with their peers. In doing so, adopted adolescents have to face the fact that they are not the same as their peers. They have a different appearance and origin, and often the timing of puberty differs as well—that is, many adopted adolescents mature earlier than do their peers.

Internationally adopted adolescents have to explore what it means to be adopted and to come from another culture, and how this knowledge can be integrated into their sense of self.

The difficulties listed above may put (internationally) adopted adolescents at increased risk of developing problem behaviors. A much-discussed question in the literature is whether adopted adolescents indeed exhibit more problem behaviors than do those who were not adopted. This is an important question because of the increasing number of international adoptions. As described above, several authors have pointed out that adolescence is often a difficult period for adopted children and their parents. However, much of their knowledge is based on clinical observations. Over the last decades, several large-scale empirical studies have examined problem behaviors of adopted adolescents in community samples, offering a more realistic perspective. However, the results of these investigations should be interpreted with caution. First, considerable heterogeneity exists within the studies. For example, the children were adopted from several different countries and cultures, the adoption procedures differed (e.g., some children stayed in institutions before placement, whereas others remained with their biological families), the ages on arrival were different, and so on. This may limit the accuracy of the conclusions. Howe, for example, found that children who were adopted after the age of 6 months had more behavior problems than did those who were adopted earlier. Second, some of the variance between adopted and nonadopted adolescents may be explained by differences in parents’ perceptions of their adopted children and biological children. Warren, for example, found a selection bias in the referral of adopted adolescents for psychiatric treatment. She reported that (domestically) adopted adolescents were more likely than nonadopted adolescents to receive psychiatric treatment, even when the level of behavior problems was controlled for.

With meta-analytic statistical techniques, the results of a large and diverse body of studies can be summarized and interpreted. Meta-analysis has been characterized as being more precise and more objective than traditional, qualitative approaches to reviewing research. Ten years ago, Wierzbicki conducted a meta-analysis of 66 published studies that compared the psychological adjustment of adoptees and nonadoptees. He found that adoptees were significantly higher in maladjustment than were nonadoptees, judging
from their representation in clinical samples. (The combined effect size \(d\) — that is, the weighted mean effect size of the 66 studies — was 1.38, where \(d\) is the standardized difference in means between the adoptees and nonadoptees. He also found more externalizing behavior problems \((d = 0.22)\) and poorer psychological functioning in adoptees of various ages than in nonadoptees (comparison of adoptees and nonadoptees on a measure of psychological functioning; \(d = 0.11\)). The combined effect size of the percentages of clinical cases and the group comparisons was \(d = 0.72\), indicating that adoptees had significantly higher levels of maladjustment. Effect sizes were significantly larger for externalizing behavior problems than for internalizing behavior problems and were larger for adolescents than for children or adults.

However, most of the studies in Wierzbicki’s meta-analysis involved nationally adopted adolescents. As mentioned above, internationally adopted adolescents may have experienced several additional disadvantageous factors. We therefore do not know whether Wierzbicki’s conclusions can be generalized to internationally adopted adolescents. Nor do we know whether the effect size would be the same for studies that involve adolescents only. In this paper, we examine problem behaviors of adolescents who were internationally adopted in infancy or early childhood, and we include some studies published subsequent to Wierzbicki’s meta-analysis.26

First, we will start with a discussion of the studies that examine problem behaviors of adolescents who were internationally adopted during infancy or early childhood. Then we will perform a meta-analysis on these studies. The purpose of this meta-analysis is to look for a general trend in the prevalence of problem behaviors in internationally adopted adolescents. On the basis of the results, we will try to determine whether internationally adopted and nonadopted adolescents differ with respect to the prevalence of problem behaviors, and if they do, in which domains these differences are manifested (i.e., internalizing vs. externalizing behavior problems). We will answer these questions for boys and girls separately. Considering the risks associated with international adoption in general, which may become more apparent during adolescence, and the results of Wierzbicki’s meta-analysis conducted on mostly nationally adopted participants,26 we expect to find substantially more problem behaviors in internationally adopted adolescents than in their nonadopted peers. In the meta-analysis we will also examine whether differences in results are related to characteristics of the studies, including both methodological variables (for example, sample selection) and characteristics of the participants (for example, mean age at adoption).
Chapter 2

Narrative review of studies

Methods

Studies were selected by means of computerized searches (PsycInfo, ERIC, Sociological Abstracts, Puis International, and Medline) on the terms “adopted adolescents,” “international adoptees,” “intercountry adoption,” and “adjustment,” as well as manual searches of the reference lists of books, articles, and a previous review of the adjustment of adoptees. Studies were included if they (1) involved an empirical investigation, (2) discussed problem behaviors (externalizing and/or internalizing) of adoptees between 12 and 20 years old, as determined by questionnaires, interviews, observations, or other measures, and (3) reported sufficient data to allow for the computation of an effect size for any difference in problem behaviors between internationally adopted adolescents and nonadopted adolescents. Because data for the Rosenwald study were incomplete, we communicated with the author, who was willing to share unpublished information. The studies originated from North America, several European countries, and Australia. The characteristics of investigations from different countries can differ (for example, in age of adoption, for which every country has its own regulations and laws). However, in our meta-analysis, moderator variables take this diversity into account and allow for tests of its influence on the combined effect size. We did not exclude unpublished studies, because the inclusion of these studies in meta-analyses is considered important to prevent publication biases from inflating the results. In addition, the total number of pertinent studies was rather small.

Two publications included the same sample examined at different stages of a longitudinal project. Because the age of the children in the second study more closely matched the years of adolescence, we decided to include only this report in the meta-analysis.

The selection procedure produced ten studies on problem behaviors in internationally adopted and nonadopted adolescents. Investigators compared either the mean scores on measures of behavior problems (we termed such data “continuous results”) or the percentage or number of children who exhibited behavior problems or scored in the clinical range (“categorical results”). Some studies presented both types of results. All samples were nonclinical, which is important because our purpose was to examine behavior problems of internationally adopted adolescents in general, rather than in a selected group of them. Besides, only one study compared the adoption sample with a nonadopted clinical sample (see below). Nine of the ten studies have been published; the tenth study was a paper based on an unpublished thesis.
Results and discussion

In Table 1, we present an overview of the basic characteristics of the studies that examined problem behaviors of internationally adopted adolescents. Andresen\(^3\) studied the behavioral and emotional adjustment of 134 12- and 13-year-old internationally adopted children living in Norway. Fifty-two percent of the children came from Korea, and 48% came from other, unspecified countries. Of the families contacted, 75% participated. The adjustment of the adopted children was assessed with Norwegian translations of the Rutter Parent\(^3\) and Rutter Teacher\(^3\) questionnaires, brief measures that differentiate between children with and without emotional or behavioral disorders. The mean scores of the total sample of adopted children on the full scale (teacher form) differed significantly from those of a matched nonadopted comparison group: adopted children had higher problem scores than did their nonadopted classmates. However, no statistically significant difference was found between the two groups regarding the number of children with scores above the cutoff point for behavioral problems of clinical significance. Thus, although the adopted children experienced somewhat more problems than did their nonadopted classmates, the problems were rarely of sufficient consequence to classify the children as maladjusted. Of the three subscales, only hyperactivity differed markedly between the two groups, with adopted children scoring significantly higher than their nonadopted classmates. The difference was statistically significant both for the total sample and for boys, but not for girls.

Bagley\(^3\) conducted a longitudinal study of the adjustment of adoptees from a variety of ethnic groups (Native Canadian Indian adoptees, white adoptees, and intercountry adoptees) and family situations. The sample of intercountry adoptees consisted of 20 adolescents aged 13–17 who were either Asian (55%) or South American (45%) and were then living in Canadian families. Of the randomly selected group of subjects, 84.5% participated. The adolescents and their parents were interviewed in their homes about child–parent relations and behavior problems (rebellion, school truancy, running away from home, substance use, delinquency, sexual acting-out) at the time of the interview. They also completed a measure of suicidal ideas and behavior.\(^3\) The intercountry adoptees did not exhibit more behavior problems or report more suicidal ideas or acts than did nonadopted Caucasian adolescents.

Berg-Kelly and Eriksson\(^9\) compared 125 13- to 18-year-old international adoptees living in Sweden with 9,204 of their classmates of these adoptees in school with respect to health, health habits, and risk behavior. Most of the adopted children came from Korea and India. Nonparticipation was estimated to be around 10% — the percentage of absenteeism from school at this age level.
The adolescents completed the Q90, a questionnaire concerning mental health (such as depressed feelings, nervousness, and suicidal thoughts), problem behaviors (such as fighting, school truancy, use of illicit drugs, and getting drunk regularly), and the acquisition of adult life styles (such as the use of alcohol and tobacco). The adopted girls, but not the adopted boys, reported suicidal thoughts, school truancy, and contact with illicit drugs significantly more often than did their nonadopted classmates.

Bogaerts and Van Aelst studied the psychosocial adjustment of 70 15- and 16-year-olds who were adopted from India and were now living in Belgium. Of the families contacted, 82% participated. The parents completed the Child Behavior Checklist (CBCL), and the adolescents completed the Youth Self-Report (YSR). The CBCL and the YSR are standardized questionnaires for quantifying a broad range of children’s problem behaviors, both externalizing (such as aggression and delinquency) and internalizing (such as anxiety/depression, somatic complaints, withdrawal) problem behaviors. Adopted youngsters scored higher on the total problem score than did their nonadopted counterparts from the general population. This was especially true for the adopted girls, who scored higher than nonadopted girls from the general population on both internalizing and externalizing problem behaviors: delinquency, withdrawal, anxiety/depression, and attention difficulties. The adopted boys scored higher than nonadopted boys from the general population on attention problems and on delinquent and aggressive behavior, which are all externalizing problem behaviors.

Cederblad and colleagues (see also: Irhammar and Cederblad) reported on the mental health of 211 adopted children who were living in Sweden and were 13 years of age or older at the time of the investigation. Most of the children were from India (36%), Thailand (15%), or Chile (15%). An additional 2% came from Ethiopia, with the remainder originating in various Asian or Latin American countries. The rate of nonresponse of the families was 19%. The adoptive mothers completed the CBCL for 133 13- to 16-year-old adopted adolescents, and the scores were compared with those of a random sample of 529 13- to 16-year-old nonadopted adolescents also living in Sweden. The mental health of the adoptees was similar to that found in the comparison group. There were no differences between different age levels. The 39 17- to 19-year-old adoptees completed the Symptom Check List, a widely used measure that contains a series of 90 items referring to expressions of psychosomatic and emotional distress. Their scores were compared with those from a random sample of 63 18- to 21-year-old nonadopted adolescents living in Sweden. The mental health of the two groups was similar, except that the adoptees had higher scores on “obsessive-compulsive” symptoms. The percentage of individuals who were severely disturbed did not differ between groups.
<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Sample (n)</th>
<th>Age of adoptees</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Adoptees</td>
<td>Nonadoptees</td>
<td>At time of study (y)* On arrival (mo)†</td>
</tr>
<tr>
<td>Andresen(^{13})</td>
<td>Norway</td>
<td>134</td>
<td>134</td>
<td>12–13 (12.5)</td>
</tr>
<tr>
<td>Bagley(^{36})</td>
<td>Canada</td>
<td>20</td>
<td>40</td>
<td>13–17 (16.3) 42 Measure of suicidal ideas and behavior</td>
</tr>
<tr>
<td>Berg-Kelly &amp; Eriksson(^{19})</td>
<td>Sweden</td>
<td>125</td>
<td>9204</td>
<td>13–18 (14.9)</td>
</tr>
<tr>
<td>Bogaerts &amp; Van Aelst(^{39})</td>
<td>Belgium</td>
<td>70</td>
<td>758</td>
<td>16–17 (16.5) 22.5 CBCL</td>
</tr>
<tr>
<td>Cederblad et al.(^{42})</td>
<td>Sweden</td>
<td>133</td>
<td>529</td>
<td>13–16 (14.9) 10 CBCL</td>
</tr>
<tr>
<td>Geerars et al.(^{20})</td>
<td>Netherlands</td>
<td>65</td>
<td>756</td>
<td>15–17 (16) 4 CBCL</td>
</tr>
<tr>
<td>Goldney et al.(^{32})</td>
<td>Australia</td>
<td>34</td>
<td>241</td>
<td>12–20 (15.2) 17 CBCL</td>
</tr>
<tr>
<td>Rosenwald(^{27})</td>
<td>Australia</td>
<td>67</td>
<td>985</td>
<td>12–16 (14) 50 CBCL</td>
</tr>
<tr>
<td>Sharma et al.(^{48})</td>
<td>USA</td>
<td>92</td>
<td>1719</td>
<td>12–18 (15) YSR</td>
</tr>
<tr>
<td>Versluis–den Bieman &amp; Verhulst(^{31})</td>
<td>Netherlands</td>
<td>1538</td>
<td>311</td>
<td>14–18 (15.3) CBCL</td>
</tr>
</tbody>
</table>

CBCL: Child Behavior Checklist; Q90: 90-item Questionnaire on adolescent health and risk behavior; SCL-90: 90-item Symptom Check List; YSR: Youth Self-Report.

*Range (estimated mean).
†If known; estimated mean.
‡The Rutter Parent Questionnaire was also examined in this study, but only the Teacher Questionnaire was used in the meta-analysis.
Geerars and colleagues\textsuperscript{20} (see also: Hoksbergen\textsuperscript{46}) reported on the adjustment of 65 15- to 20-year-olds who were adopted from Thailand and were then living in the Netherlands. Of the individuals contacted, 24\% declined to participate. The parents and the adoptees were asked to complete the CBCL and the YSR, respectively, and their scores were compared with those of a group of 756 12- to 16-year-old non-adopted Dutch children. Adopted boys and girls both scored higher on the total problem score of the CBCL than did the nonadopted adolescents, but the difference was statistically significant only for the girls. Adopted girls scored significantly higher for anxious/obsessive, depressed/withdrawn, schizoid, and delinquent behaviors, while boys scored significantly higher only on delinquent and aggressive behaviors.

Goldney and colleagues\textsuperscript{32} compared the prevalence of emotional and behavioral problems in 34 adolescent adoptees from Indonesia living in South Australian families with that of 233 nonadopted adolescents. Eighty-five percent of the families contacted participated in the study. The adoptees completed the YSR, and their adoptive mothers completed the CBCL. No significant differences were found between the scores of the adopted and the nonadopted groups. This was true for both boys and girls. The authors also compared the scores of the adoptees with the scores of adolescents who were referred to mental health clinics. The adoptees had significantly fewer problems than did the adolescents from the clinic population. Because none of the other studies compared the adoption sample with a sample of clinically referred nonadopted adolescents, it was not possible to include this comparison in the meta-analyses.

Rosenwald\textsuperscript{27} used parental reports on the CBCL to assess the well-being of 67 internationally adopted adolescents aged 12–16 who were living in Western Australia. Seventy-five percent of the children came from Korea, 7\% or less each from India, Sri Lanka, Hong Kong, Mauritius, Philippines, and Fiji, and 5\% from all other countries combined. Eighty-six percent of the families that were contacted participated. The author compared the adopted children with 985 nonadopted adolescents from the Western Australian Child Health Survey.\textsuperscript{47} Twenty-three percent of the adopted boys and 20\% of the nonadopted boys showed significant levels of behavior problems. Among girls, the figures were 17\% and 15\%, respectively.

Sharma and coworkers\textsuperscript{48} used the YSR to compare 92 internationally adopted Asian-American (primarily Korean-American) adolescents with 1719 nonadopted adolescents aged 11–18. Of the families contacted, 54\% participated. Adopted boys showed poorer adjustment than did the boys in the comparison group on the Self-Destruct scale ($d = 0.36$), and adopted girls showed a higher level of adjustment than did the girls in the comparison group on the Social Problems scale ($d = 0.36$) and the Withdrawn scale ($d = 0.42$). The effect sizes for the Externalizing, Internalizing, and Total Problem scales were very small. The authors also compared adopted adolescents with adolescents
Problem Behavior of Adopted Adolescents

born to their adoptive parents and raised in the same families, but because not all of the adoptions were international, we chose to use only the sample of internationally adopted Asian-American adolescents in the meta-analysis.

Versluis–den Bieman and Verhulst conducted a large cross-sectional study of the prevalence of self- and parent-reported problems in a sample of 1538 14- to 18-year-old international adoptees who were living in the Netherlands. The youngsters came from Korea (33.9%), Colombia (14.3%), India (9.9%), Indonesia (7.7%), Bangladesh (6.8%), Lebanon (5.1%), Austria (5.0%), and other European (3.3%) and non-European (14.0%) countries. Usable parent information was obtained on 74% of the children. The parents and adopted adolescents completed the CBCL and the YSR, respectively. Both parent reports and self-reports showed significantly higher problem scores for adopted boys and girls than for adolescents from the general population. According to self-reports, behavior could be regarded as deviant in 22% of the adopted boys and 18% of the adopted girls compared with about 10% of the nonadopted children. The difference between adopted and nonadopted boys was larger in the parent reports, however. According to these reports, behavior of 29% of the adopted boys versus 9% of the boys from the general population could be regarded as deviant. The largest difference between the groups was for “Delinquent Behavior syndrome”: the proportion of boys scored in the deviant range was ten times as high among the adoptees as among the nonadoptees. For girls, parent reports and self-reports were quite similar, with 17% of the adopted girls and 10% of the girls from the general population manifesting deviant behavior.

Meta-analysis

Methods

The selection procedure yielded ten studies on the difference in problem behaviors between internationally adopted and nonadopted adolescents. In Cederblad and colleagues’ investigation the sample was split in two subsamples. Mothers completed the CBCL/4-18 for adolescents aged 13–16, and subjects over 16 years old completed the Symptom Check List. Therefore, the number of samples in the meta-analysis was 11, which included 17,057 adolescents (2317 internationally adopted, 14,740 nonadopted). Because Bogaerts and Van Aelst, Geerars and colleagues, and Versluis–den Bieman and Verhulst all used the same comparison group (the group of adolescents used to derive the Dutch CBCL-norms from), we split up the comparison group in the meta-analysis (see Table 1; Bogaerts and Van Aelst, 2023 comparison subjects; Geerars et al., 2023; Versluis–den Bieman and Verhulst, 2023).
311). (Versluis–den Bieman and Verhulst, in fact, randomly selected 311 adolescents from the total comparison group.) This resulted in a total sample size for the meta-analysis of 16,662 adolescents, of whom 2317 were internationally adopted and 14,345 were not adopted.

Bogaerts and Van Aelst, Goldney and colleagues, and Versluis–den Bieman and Verhulst used both the CBCL and the YSR, but we chose to use only the CBCL in the meta-analysis, because the YSR is a self-report instrument that may produce somewhat less reliable results. Moreover, the CBCL is used more often in the other studies, which makes comparisons among the studies more reliable. For investigations that only used self-reports, we employed these self-reports in the meta-analysis.

In primary-level studies, the unit of analysis is the participant; in a meta-analysis of several primary-level studies, the unit of analysis is the outcome of those studies. The outcomes found in the relevant studies were first transformed into a common meta-analytic indicator for effect size, the standardized difference between the means of two groups (Cohen’s $d$). For studies that reported test statistics for comparisons of internationally adopted adolescents and nonadopted adolescents, these statistics ($t$, which reflects the difference between the means of two independent samples; $F$, which reflects the difference among means of more than two populations simultaneously; $\chi^2$, which shows the fit of observed frequencies with expected frequencies; or the one-directional $p$ value, which indicates the probability of obtaining a value as extreme as that found in the sample when the null hypothesis is true) were transformed into $d$ using Mullen’s computer program Advanced Basic Meta-Analysis. If a study reported only “a significant effect” or “no significant effect,” we applied the usual conservative estimation procedures ($p = 0.05$ and $p = 0.50$, respectively). A few studies did not report $t$, $F$, $\chi^2$, or the one-directional $p$ value, giving only the number or percentage of adolescents who exhibited behavior problems or scored in the clinical range. For these studies, $\chi^2$ values were first computed using the statistical programs Fisher 3.0 and Multinom and then transformed into $d$. If possible, separate effect sizes were calculated for boys and girls, and for internalizing (e.g., anxiety, depression, somatic complaints, withdrawal, suicidal ideas/thoughts, nervousness) and externalizing (aggression, delinquency, hyperactivity, antisocial behavior, physical fights) behavior problems. When the scores for internalizing and externalizing behavior problems were not reported, the effect sizes for them were calculated from the scores of the subscales that were usually included in these overall problem scales. In this way, each study contributed, at most, one effect size for each type of behavior problem, and no sample was included more than once in the meta-analysis.

Combined effect sizes, which are the weighted mean effect sizes of all studies included, were determined with the computer program Comprehensive
Meta-Analysis. Combined effect sizes were calculated separately for the total group, boys, girls, total behavior problems, externalizing behavior problems, and internalizing behavior problems. Moreover, we distinguished between continuous results (means) and categorical results (cutoff scores or percentage/number of adolescents exhibiting behavior problems). We also calculated combined effect sizes for the total set of studies—that is, for continuous and categorical results together. If one study reported both types of results, we chose the more accurate or more precise type for these analyses. A homogeneity test was performed to determine to what extent effect sizes were constant across studies.

The following predictor variables were included in the meta-analyses:

1. Mean age of the adoptees at the time of the study (early adolescence, 12 or 13 years; mid-adolescence, 14–16 years; late adolescence, 17 or 18 years),
2. Mean age on arrival (<12 months, 12–24 months, >24 months)
3. Measure (whether or not the study used the CBCL)
4. Self/other ratings (self-ratings, parent ratings, teacher ratings)
5. Sex (boys or girls)
6. Socioeconomic status of the sample (lower class, middle class, upper-middle class)
7. Year of publication
8. Country where the study was conducted
9. Continent where the study was conducted (North America, Europe, Australia)
10. Sample selection (comparison groups recruited randomly or age- and sex-matched with the adoption sample)
11. Attrition (<22.1% (mean attrition of all studies) or >22.1%)
12. Type of results (continuous or categorical results).

When a study did not conform exactly to the categories listed above, it was assigned to the category that most closely matched it. The mean intercoder reliability of the predictors was \( r = 0.92 \) (range, 0.73–1.00). Analyses of variance were conducted to determine whether a predictor or moderator variable was associated with the effect sizes.

Results and discussion

In Table 2, we present an overview of the types of results (for example, total number of behavior problems, externalizing problems, internalizing problems) reported in studies of internationally adopted adolescents. Table 3
Chapter 2

summarizes the relevant meta-analytic data. The data are presented separately for externalizing, internalizing, and total behavior problems, and for boys and girls. We also distinguished between the outcomes obtained from continuous results (means) and those obtained from categorical results and show outcomes separately for the two (see Methods).

Table 2

<table>
<thead>
<tr>
<th>Study</th>
<th>Total number of problems</th>
<th>Separate results: externalizing and internalizing behaviors</th>
<th>Separate results: boys and girls</th>
<th>Mean scores</th>
<th>Cutoff scores</th>
<th>Percentage or number of children with behavior problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andresen</td>
<td>x</td>
<td>x*</td>
<td>x</td>
<td>x</td>
<td>x†</td>
<td>x</td>
</tr>
<tr>
<td>Bagley</td>
<td>x</td>
<td>x‡</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Berg-Kelly &amp; Eriksson</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Bogaerts &amp; Van Aelst</td>
<td>x</td>
<td>x*</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Cederblad et al.</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Geerars et al.</td>
<td>x</td>
<td>x*</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Goldney et al.</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Rosenwald</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Sharma et al.</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Versluis–den Bieman &amp; Verhulst</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

*We computed the scores for externalizing and internalizing behavior problems from the scores of the relevant subscales.
†We used the cutoff scores as the more-accurate results in the meta-analysis. Because the cutoff scores did not distinguish between internalizing and externalizing problems, we used the means for these problems in the meta-analysis.
‡For the means, it was not possible to calculate exact t- or p-values because we did not have exact standard deviations. Therefore, we used the cutoff scores as the more-accurate results in the meta-analysis. We compared the cutoff scores with Dutch CBCL norms.
§We compared the cutoff scores with Dutch CBCL norms, but we did not use these scores in the meta-analysis because of their extremely high effect sizes. These results would have been outliers in the meta-analysis.
Sharma et al. reported standardized effect sizes of the scores.
<table>
<thead>
<tr>
<th>Type of results</th>
<th>Group</th>
<th>k</th>
<th>Na</th>
<th>Nc</th>
<th>Ntot</th>
<th>d</th>
<th>CI</th>
<th>p</th>
<th>Q</th>
<th>Q</th>
<th>Q</th>
<th>Q</th>
<th>Q</th>
<th>Q</th>
<th>Q</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>Boys</td>
<td>7</td>
<td>940</td>
<td>6495</td>
<td>7435</td>
<td>0.13</td>
<td>0.01~0.24</td>
<td>0.03</td>
<td>4.63</td>
<td>0.53</td>
<td>1.35</td>
<td>0.10</td>
<td>1.23</td>
<td>3.37</td>
<td>2.92</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>7</td>
<td>1118</td>
<td>6234</td>
<td>7352</td>
<td>0.13</td>
<td>0.03~0.23</td>
<td>0.01</td>
<td>2.68</td>
<td>0.69</td>
<td>0.73</td>
<td>1.06</td>
<td>0.98</td>
<td>0.07</td>
<td>0.17</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>10</td>
<td>2250</td>
<td>13360</td>
<td>15610</td>
<td>0.11</td>
<td>0.04~0.18</td>
<td>0.00</td>
<td>5.86</td>
<td>0.71</td>
<td>2.76</td>
<td>0.01</td>
<td>1.18</td>
<td>1.45</td>
<td>4.41*</td>
<td>6.25*</td>
</tr>
<tr>
<td>Externalizing</td>
<td>Boys</td>
<td>7</td>
<td>940</td>
<td>6495</td>
<td>7435</td>
<td>0.07</td>
<td>-0.04~0.19</td>
<td>0.21</td>
<td>5.73</td>
<td>0.19</td>
<td>4.13</td>
<td>0.84</td>
<td>0.48</td>
<td>0.22</td>
<td>3.07</td>
<td>3.48</td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>7</td>
<td>1118</td>
<td>6234</td>
<td>7352</td>
<td>0.05</td>
<td>-0.05~0.15</td>
<td>0.34</td>
<td>7.57</td>
<td>0.09</td>
<td>1.93</td>
<td>0.79</td>
<td>2.81</td>
<td>1.28</td>
<td>2.86</td>
<td>2.57*</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>10</td>
<td>2250</td>
<td>13360</td>
<td>15610</td>
<td>0.05</td>
<td>-0.01~0.12</td>
<td>0.12</td>
<td>11.43</td>
<td>0.22</td>
<td>6.26</td>
<td>1.33</td>
<td>1.33</td>
<td>0.34</td>
<td>4.99*</td>
<td>5.98*</td>
</tr>
<tr>
<td>Internalizing</td>
<td>Boys</td>
<td>7</td>
<td>940</td>
<td>6495</td>
<td>7435</td>
<td>0.07</td>
<td>-0.04~0.19</td>
<td>0.21</td>
<td>5.73</td>
<td>0.19</td>
<td>4.13</td>
<td>0.84</td>
<td>0.48</td>
<td>0.22</td>
<td>3.07</td>
<td>3.48</td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>7</td>
<td>1118</td>
<td>6234</td>
<td>7352</td>
<td>0.05</td>
<td>-0.05~0.15</td>
<td>0.34</td>
<td>7.57</td>
<td>0.09</td>
<td>1.93</td>
<td>0.79</td>
<td>2.81</td>
<td>1.28</td>
<td>2.86</td>
<td>2.57*</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>10</td>
<td>2250</td>
<td>13360</td>
<td>15610</td>
<td>0.05</td>
<td>-0.01~0.12</td>
<td>0.12</td>
<td>11.43</td>
<td>0.22</td>
<td>6.26</td>
<td>1.33</td>
<td>1.33</td>
<td>0.34</td>
<td>4.99*</td>
<td>5.98*</td>
</tr>
<tr>
<td>Total</td>
<td>Boys</td>
<td>8</td>
<td>167</td>
<td>1263</td>
<td>1430</td>
<td>0.08</td>
<td>-0.10~0.26</td>
<td>0.39</td>
<td>4.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>8</td>
<td>172</td>
<td>1279</td>
<td>1511</td>
<td>0.10</td>
<td>0.01~0.20</td>
<td>0.03</td>
<td>3.60</td>
<td>0.08</td>
<td>0.56</td>
<td>0.59</td>
<td>2.01</td>
<td>1.49</td>
<td>1.18</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>11</td>
<td>337</td>
<td>2541</td>
<td>2880</td>
<td>0.08</td>
<td>0.01~0.14</td>
<td>0.02</td>
<td>9.52</td>
<td>0.57</td>
<td>5.62</td>
<td>0.05</td>
<td>1.42</td>
<td>1.03</td>
<td>6.74**</td>
<td>5.49*</td>
</tr>
<tr>
<td>Continuous</td>
<td>Boys</td>
<td>5</td>
<td>167</td>
<td>1263</td>
<td>1430</td>
<td>0.08</td>
<td>-0.10~0.26</td>
<td>0.39</td>
<td>4.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>5</td>
<td>228</td>
<td>1278</td>
<td>1506</td>
<td>0.16</td>
<td>0.01~0.31</td>
<td>0.04</td>
<td>1.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>7</td>
<td>567</td>
<td>3132</td>
<td>3699</td>
<td>0.08</td>
<td>-0.02~0.17</td>
<td>0.13</td>
<td>3.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internalizing</td>
<td>Boys</td>
<td>5</td>
<td>167</td>
<td>1263</td>
<td>1430</td>
<td>0.08</td>
<td>-0.10~0.26</td>
<td>0.37</td>
<td>1.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>5</td>
<td>228</td>
<td>1278</td>
<td>1506</td>
<td>0.06</td>
<td>-0.09~0.22</td>
<td>0.41</td>
<td>3.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>7</td>
<td>567</td>
<td>3132</td>
<td>3699</td>
<td>0.05</td>
<td>-0.05~0.15</td>
<td>0.31</td>
<td>2.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Boys</td>
<td>5</td>
<td>167</td>
<td>1263</td>
<td>1430</td>
<td>0.05</td>
<td>-0.13~0.23</td>
<td>0.56</td>
<td>1.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>5</td>
<td>228</td>
<td>1278</td>
<td>1506</td>
<td>0.10</td>
<td>-0.06~0.25</td>
<td>0.22</td>
<td>2.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>6</td>
<td>333</td>
<td>2891</td>
<td>3424</td>
<td>0.09</td>
<td>-0.01~0.19</td>
<td>0.07</td>
<td>4.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of results</td>
<td>Group</td>
<td>k</td>
<td>Na</td>
<td>Nc</td>
<td>Ntot</td>
<td>d</td>
<td>CI</td>
<td>p</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
</tr>
<tr>
<td>----------------</td>
<td>-------</td>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Categorical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Externalizing</td>
<td>Boys</td>
<td>3</td>
<td>806</td>
<td>5341</td>
<td>6147</td>
<td>0.17</td>
<td>0.03~0.31</td>
<td>0.02</td>
<td>0.03~0.31</td>
<td>0.02</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>3</td>
<td>927</td>
<td>5070</td>
<td>5997</td>
<td>0.13</td>
<td>0.01~0.26</td>
<td>0.04</td>
<td>0.13</td>
<td>0.01~0.26</td>
<td>0.04</td>
<td>1.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>4</td>
<td>1753</td>
<td>10451</td>
<td>12204</td>
<td>0.15</td>
<td>0.06~0.24</td>
<td>0.00</td>
<td>0.15</td>
<td>0.06~0.24</td>
<td>0.00</td>
<td>1.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internalizing</td>
<td>Boys</td>
<td>3</td>
<td>806</td>
<td>5341</td>
<td>6147</td>
<td>0.10</td>
<td>-0.05~0.24</td>
<td>0.18</td>
<td>5.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>3</td>
<td>927</td>
<td>5070</td>
<td>5997</td>
<td>0.06</td>
<td>-0.06~0.19</td>
<td>0.33</td>
<td>4.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>4</td>
<td>1753</td>
<td>10451</td>
<td>12204</td>
<td>0.08</td>
<td>-0.02~0.17</td>
<td>0.11</td>
<td>10.16*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Boys</td>
<td>5</td>
<td>883</td>
<td>5885</td>
<td>6768</td>
<td>0.09</td>
<td>-0.04~0.22</td>
<td>0.16</td>
<td>10.90*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Girls</td>
<td>5</td>
<td>1051</td>
<td>5645</td>
<td>6696</td>
<td>0.11</td>
<td>0.00~0.22</td>
<td>0.04</td>
<td>1.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>6</td>
<td>1954</td>
<td>11570</td>
<td>13524</td>
<td>0.10</td>
<td>0.02~0.18</td>
<td>0.01</td>
<td>7.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Age, (estimated) mean age of the adoption group; Age a, (estimated) mean age at arrival of the adoption group; Att, attrition; Cont, continent; CI, confidence interval (lower and upper limit of the 95% confidence interval); d, Cohen's d; k, number of studies; Meas, type of measurement; Na, number of children in the adoption group; Nc, number of children in the comparison group; Ntot, total number of children included in the study; p, probability; Q, test of heterogeneity; Rep, type of report; Res, type of results; Samp, sample selection.

* p < 0.05; ** p < 0.01.
†Because of the homogeneity of the results, we used fixed effects instead of random effects.
Combined effect sizes

For the complete set of studies, the combined effect size for total behavior problems of the entire group (boys and girls together) was $d = 0.08$ (one-tailed $p = 0.02$; see Table 3). According to Cohen’s criteria for weak ($d = 0.20$), medium ($d = 0.50$), and strong ($d = 0.80$) effects, this effect was weak. The combined effect size for externalizing behavior problems of the entire group was $d = 0.11$ ($p < 0.01$), which was somewhat larger than the combined effect size for total behavior problems but may still be considered small. The combined effect size for internalizing behavior problems was even smaller and did not reach significance ($d = 0.05$; $p = 0.12$). For boys, the combined effect size for total behavior problems did not reach significance ($d = 0.07$, $p = 0.22$). The combined effect sizes for externalizing and internalizing behavior problems for boys were $d = 0.13$ ($p = 0.03$) and $d = 0.07$ ($p = 0.21$), respectively. Compared with nonadopted adolescent boys, internationally adopted boys exhibited more externalizing behavior problems, but this difference was small; they did not show more internalizing behavior problems. For girls, the combined effect size for total behavior problems was $d = 0.10$ ($p = 0.03$). The combined effect sizes for externalizing and internalizing behavior problems for girls were $d = 0.13$ ($p = 0.01$) and $d = 0.05$ ($p = 0.34$), respectively. Summarizing, internationally adopted adolescents exhibited more total behavior problems than did nonadopted adolescents. If we look at boys and girls separately, only the girls exhibited more total behavior problems. Differences between the adopted and nonadopted adolescents were, in all groups (entire group, boys, and girls), revealed in externalizing but not in internalizing behavior problems. All differences were small.

No outlying effect sizes were identified in the set of studies on the basis of standardized scores ($z$-values) larger than 3.26 or smaller than -3.26 ($p < 0.001$). To check the contribution of the largest study (the study by Versluis–den Bieman and Verhulst, which showed the largest standardized $z$-value [1.95] for total behavior problems) on the combined effect size, we computed the overall effect sizes without this investigation. The small, but significant, combined effect sizes for all studies were no longer apparent when we excluded this particular one. Thus, without the Versluis–den Bieman and Verhulst study, adopted and nonadopted adolescents did not differ significantly with respect to the prevalence of behavior problems.

We also distinguished between continuous results and categorical results. The combined effect sizes of the continuous results were all very small (see Table 3). Only the effect size of externalizing behavior problems of girls was significant. The effect sizes of the categorical results were also small, albeit somewhat larger than those of the continuous results. Effect sizes were significant for externalizing behavior problems of all groups (boys, girls, and the entire group) and for total behavior problems of the girls and the entire group.
So, the effect sizes were larger when the percentage or number of internationally adopted adolescents who exhibited behavior problems or scored in the clinical range were compared with the percentage or number of nonadopted adolescents who did so. The effect sizes were smaller when the mean scores of internationally adopted and nonadopted adolescents were compared. This difference was again no longer apparent when we excluded the Versluis–den Bieman and Verhulst study, which compared the percentage of internationally adopted adolescents scoring in the clinical range with the corresponding percentage of nonadopted adolescents.

**Homogeneity**

A homogeneity test with the Q-statistic was performed to determine to what extent effect sizes were constant across studies. All sets of study results were homogeneous (see Table 3) except for two: the effect sizes for the categorical results of total behavior problems of boys ($Q = 10.90; p < 0.05$) and of internalizing behavior problems of the entire group ($Q = 10.16; p < 0.05$). Aside from these two heterogeneous outcomes, the combined or average effect sizes were therefore an adequate representation of the sets of study outcomes. Without the Versluis–den Bieman and Verhulst investigation, all sets of study results were homogeneous.

**Predictor tests**

Analyses of variance were conducted to determine whether the predictor variables had influence on the effect sizes. Sample selection (whether participants in the comparison groups were recruited randomly or were age- and sex-matched with the adoption sample) contributed significantly to the predictability of the combined effect sizes of the total group (both boys and girls) on externalizing ($Q = 4.41; p < 0.05$), internalizing ($Q = 4.99; p < 0.05$) and total behavior problems ($Q = 6.74; p < 0.01$), as well as to the predictability of the effect size for boys on total behavior problems ($Q = 6.90; p < 0.01$). Studies that randomly recruited participants in the comparison groups showed lower effect sizes than did those that recruited age- and sex-matched participants. Although not significant, this trend was also seen for the other outcome measures.

Mean age on arrival was correlated with the difference in the effect size of total behavior problems of adolescent boys ($Q = 8.97; p < 0.05$). For boys with a mean age of arrival between 12 and 24 months, the combined effect size was larger and negative ($d = -0.28$; i.e., adopted boys showed fewer behavior problems than did nonadopted ones); the combined effect sizes for boys with a mean arrival age between 0 and 12 months ($d = 0.02$) and boys who were older than 24 months when they were adopted ($d = 0.03$) were smaller and positive. For the other outcomes, this trend was not seen.
Attrition contributed significantly to the predictability of the combined effect sizes of the entire group (both boys and girls) on externalizing \((Q = 6.25; p < 0.05)\), internalizing \((Q = 5.98; p < 0.05)\), and total behavior problems \((Q = 5.49; p < 0.05)\), for the boys on total behavior problems \((Q = 5.91; p < 0.05)\), and for the girls on internalizing behavior problems \((Q = 2.57, p < 0.05)\). Studies that had an attrition rate lower than 22.1% (the mean for all studies) showed smaller effect sizes than did those with a rate of 22.1% or more. None of the other predictor or moderator variables contributed significantly to the predictability of the effect sizes.

General discussion and conclusions

We investigated the prevalence of problem behaviors in samples of adolescents who were adopted as infants or young children, and we explored in which domains these problems are manifested. The narrative review shows that two studies on problem behaviors of internationally adopted adolescents (those by Versluis–den Bieman and Verhulst\(^{31}\) and Bogaerts and Van Aelst\(^{39}\)) reported significantly more behavior problems in such children than in nonadopted youngsters. Two studies (those by Geerars et al.\(^{20}\) and Berg-Kelly and Eriksson\(^{39}\)) found significantly more behavior problems only in adopted girls, and five (Bagley,\(^{36}\) Goldney et al.,\(^{32}\) Cederblad et al.,\(^{42}\) Sharma et al.,\(^{48}\) and Rosenwald\(^{27}\)) showed no more behavior problems in adopted adolescents, either boys or girls, than in their nonadopted peers. Andresen\(^{33}\) found a significant difference in the mean scores on behavior-problem scales between the adopted and nonadopted adolescents, but not in the number of adolescents with scores above the cutoff point. In boys, the differences were apparent mainly in externalizing behaviors (e.g., hyperactivity, aggression, and delinquent behaviors). In girls, both externalizing and internalizing behavior problems (anxiety/depression, withdrawal, schizoid and delinquent behaviors) were responsible for the differences.

Our meta-analysis showed that internationally adopted adolescents appeared to exhibit a slightly higher number of total behavior problems than did nonadopted adolescents. Considering boys and girls separately, we found that in fact only girls displayed more total behavior problems. Furthermore, in all comparisons (entire group, boys, and girls), the differences between the adopted and nonadopted adolescents were expressed in externalizing but not in internalizing behavior problems. According to Cohen’s\(^{54}\) criteria for weak, medium, and strong effect sizes, all effects were weak, with \(d\) ranging from >0.05 to <0.13.

Our results are consistent with those from Wierzbicki’s group comparisons\(^{26}\) between (mostly nationally adopted) adoptees and nonadoptees.
on a measure of psychological functioning \((d = 0.11)\). Wierzbicki also found higher maladjustment in adoptees compared with nonadoptees. In addition, he reported a stronger combined effect size for externalizing behavior problems than for internalizing ones. Wierzbicki’s reported effect sizes for mostly nationally adopted adolescents do not differ appreciably from ours. He computed effect sizes of \(d = 0.11, 0.22,\) and \(-0.01\) for total problems, externalizing behavior problems, and internalizing behavior problems (group comparisons), respectively. Accordingly, despite the many additional disadvantageous factors that internationally adopted adolescents may experience, they do not seem to exhibit more behavior problems than do mostly nationally adopted children. This finding is in line with empirical investigations comparing internationally adopted children of different races with same-race adopted children.\(^{36,56}\)

The differences in behavior problems between the adopted and nonadopted adolescents are somewhat larger when we compare the percentage or number of internationally adopted adolescents who exhibit behavior problems or score in the clinical range with the corresponding percentage or number of nonadopted adolescents (categorical comparisons; effect sizes ranging from \(d > 0.06\) to \(d < 0.17\)). Wierzbicki\(^{26}\) also found that the mean effect size for comparisons between adoptees and nonadoptees on the percentage of participants showing emotional and/or behavioral problems in the clinical range \((d = 1.38)\) was larger than the mean effect size for continuous outcomes \((d = 0.11)\). However, our effect sizes for categorical results are not exactly comparable with Wierzbicki’s effect size for clinical problems. For the computation of combined effect sizes, we included studies that reported cutoff scores or the percentage or number of adolescents exhibiting behavior problems. Wierzbicki included studies that reported the percentage of adoptees in a clinical population, which resulted in a much larger effect size. So, differences in behavior problems between adopted and nonadopted children appear to be small, but a larger number of internationally adopted adolescents show behavioral problems in the clinical range.\(^{57}\) In fact, these “clinical cases” may be responsible for the elevated rate of behavior problems in the entire group of internationally adopted adolescents.\(^{30,58}\) Nevertheless, according to Cohen’s criteria,\(^{34}\) even our findings for the categorical outcomes appear to be weak at best.

We also examined the influence of study and sample characteristics on the outcomes of the studies. Sample selection appeared to be an important predictor: studies in which the comparison group was recruited randomly showed smaller effect sizes than did those in which recruitment involved matching. Matching provides a better sample for comparison, but in our case its influence may be due to the small number of studies using matched samples and to the strong influence of the large Versluis–den Bieman and Verhulst\(^{31}\) study, which employed them. The rate of attrition also appeared to be an
Problem Behavior of Adopted Adolescents

important predictor: studies with low attrition rates showed smaller effect sizes than did those with higher attrition. The use of volunteer subjects might affect the results if the most disturbed adolescents and families drop out. However, in our meta-analyses the studies with higher attrition rates showed larger effect sizes (the adopted children showed more behavior problems), so this potential bias was not apparent.

The other study and sample characteristics, such as mean age of the adoptees on arrival, measure, self/other ratings, sex, socioeconomic status, country of the study, and continent of the study, did not show significant association with outcomes. Howe\textsuperscript{23} investigated studies examining the development of children who were adopted before the age of 6 months and those who were adopted later. He found that children who were adopted after the age of 6 months had more behavior problems than did those who were adopted earlier. However, our meta-analysis did not show such an effect. An explanation may be found in the heterogeneity of the ages on arrival in most of the investigations in our meta-analyses. In most studies, there was a broad range of ages on arrival (Goldney and colleagues,\textsuperscript{32} for example, reported a range of 1 month to 7.5 years). Because of this heterogeneity, it is possible that the mean age on arrival and the category (<12 months, 12-24 months, >24 months) in which the study was placed in the current meta-analysis were not perfectly mirroring the broad range of ages in the specific sample. Also, we were able to compute a mean age on arrival for only seven of the studies. However, Howe’s study of children who were placed for adoption at different ages and who had varying pre-placement experiences\textsuperscript{59} showed that the risk factor for problem behaviors is not simply late age at placement, but rather a combination of late age at placement and poor pre-placement care from an early age. Only children who were placed at a late age and experienced poor pre-placement care from a very young age were at significant risk of showing problem behaviors during adolescence. Children who were placed at a later age but who had relatively good care during the first year or two of life showed fewer behavior problems.

Interestingly, the results of our meta-analysis changed when Versluis-den Bieman and Verhulst’s study\textsuperscript{31} was excluded from our analyses. In fact, the weak but significant increase in behavior problems in adoptees that we found when all studies were included was no longer apparent when we omitted this one. This finding may well be due to the researchers’ large sample size, which comprised almost two-thirds of the adoptees in our meta-analysis. Another explanation may be found in specific sample characteristics of Versluis-den Bieman and Verhulst’s study. The children in this study experienced relatively serious environmental adversity before adoption, placing them at an increased risk for later behavior problems. In several reports, these authors\textsuperscript{31,60,61} have indeed stressed that a large proportion of the adopted children involved in their
longitudinal study had been subjected to extremely adverse conditions, and that such adversities were associated with the prevalence of behavior problems during adolescence. We suggest that international adoption, as such, may not exert a negative effect on the prevalence and rate of adolescents’ behavior problems unless the children have been faced with serious adverse pre-adoption circumstances. However, we do not know whether this is a characteristic specific to Versluis–den Bieman and Verhulst’s study, since the other investigations in the meta-analysis either did not mention such circumstances or provided no information at all about the adopted child’s pre-placement history. Many of the children available for adoption do indeed have negative early experiences because of nonoptimal conditions in their biological family and/or country of origin.

The results of our review and meta-analysis should be interpreted with caution for several reasons. First, considerable heterogeneity exists within the population of individuals who are adopted. Differences are seen not only in the personal characteristics of the adopted children, but also in the circumstances before their adoption. For example, in most of the studies included in our meta-analysis, the children were adopted from various different countries and cultures, the adoption procedures differed (e.g., some children stayed in institutions before placement, while others lived with their biological family or in a foster home), and the ages on arrival were different. Within the population of adopted children, there may be subgroups of youngsters who are at risk for the development of adjustment problems. Therefore, until additional research on adoption and adjustment in various subgroups is completed, clinicians and policy makers must be careful to avoid generalizing this risk to the entire population of adopted individuals. Nevertheless, effect sizes were homogeneous across the studies, so the computation of combined effect sizes is meaningful.

Second, the meta-analytic evidence presented in this paper is comparative, and the causal nature of the association between international adoption and problem behaviors has not been established. We cannot eliminate the possibility that for some internationally adopted children, problem behaviors (or a genetic disposition for such behaviors) may have been the reason why the children were given up for adoption. In this case, problem behaviors are a cause rather than a consequence of the adoption.

Finally, some of the differences between adopted and nonadopted adolescents may be explained by different perceptions expressed by parents of adopted children and parents of biological children. Adoptive parents, due to higher socioeconomic status and education, may notice or identify problems sooner. In addition, adoptive parents have already had contact with social service agencies, and they know (and perhaps sometimes expect) that adoption
can bring problems, which may also make them more sensitive to behavioral difficulties.

Although many studies show more behavior problems in adopted adolescents, the elevated problem rate may be caused by a small minority of adopted adolescents. We found that the largest — although still rather small — differences between adopted and nonadopted adolescents emerged when differences in percentage of children scoring in the clinical range were examined. Differences between groups of adopted and nonadopted children may reflect the presence of a small number of severely disturbed children, possibly with extremely adverse pre-placement histories. The majority of the adopted adolescents are well adjusted and do not show more problem behaviors than do their nonadopted peers. And like most of the problems experienced in adolescence, adoptees’ problems may well diminish or disappear in early adulthood. The finding that the majority of internationally adopted adolescents are well adjusted is important from the perspective of adoption policy, too. Contrary to the expectation that, because of the potential additional risks, internationally, transracially adopted adolescents will present many behavior problems, the internationally adopted adolescents in our meta-analysis appeared to be relatively problem-free. Moreover, the rates of problem behavior seen in these adolescents were comparable to those found in national, mostly same-race adoptions. Thus, no additional risks were found for international, transracial adoptions. This positive finding should contribute to the longstanding debate on the supposed negative outcomes of international and transracial adoptions.

Finally, it should be stressed that adoption itself is not a risk factor in the adjustment of children. Adoption may even be considered a protective factor. Researchers have compared adopted children with children who were born illegitimately but remained with their biological mothers, with children who were placed in an institution, and with children who were given up for adoption but later restored to their biological parents. In these investigations the adopted children appeared to be better adjusted than were the children against whom they were compared. Moreover, children who experienced severe deprivation in an institutional setting showed substantial developmental recovery after adoption placement. Therefore, adoption may often be the best solution for a child who would otherwise be raised in an institution under miserable circumstances, despite the problems that some adopted children experience in adolescence. Nevertheless, parents, social workers, and other professionals should be aware of the disadvantages that many adopted children have experienced, and of the influence that such disadvantages may exert on the adopted children’s behavior, so that they can understand them better, interpret their behavior problems in an adequate way, and foster a warm and
Chapter 2

stable relationship\textsuperscript{78} that supports the adoptees’ development through the turbulent phase of adolescence into adulthood.

References


15. Brodzinsky DM. Adjustment to adoption: a psychosocial perspective.  


17. Hoopes JL. Adoption and identity formation. In: Brodzinsky DM,  
Schechter MD, eds. The psychology of adoption. New York: Oxford  

18. Grotevant HD, McRoy RG. Adopted adolescents in residential  
treatment: the role of the family. In: Brodzinsky DM, Schechter MD,  
eds. The psychology of adoption. New York: Oxford University Press,  

mid-adolescence as indicated by aspects of health and risk taking—a  

20. Geerars H, Hoksbergen R, Rooda J. Geadopteerden op weg naar  
volwassenheid: de integratie van 68 Thaise jongeren in de Nederlandse  

21. Oostdijk W, Yap YN, Slijper FME, Wit JM, Drop SLS. Puberteit en  
eindlengte bij uit het buitenland geadopteerde kinderen. Tijdschr  

22. Proos LA, Hofvander Y, Tuvelo T. Menarcheal age and growth  
pattern of Indian girls adopted in Sweden, I: Menarcheal age. Acta  

23. Howe D. Patterns of adoption: nature, nurture, and psychosocial  

24. Warren SB. Lower threshold for referral for psychiatric treatment for  
7.


27. Rosenwald T. Intercountry adoptive families in Western Australia: the  
well-being of their four to sixteen year-old adoptees. Paper presented in  
Utrecht, The Netherlands at the Adoption Center, Utrecht University,  

28. Mullen B. Advanced BASIC meta-analysis. Hillsdale, New Jersey:  


