LONG-TERM OUTCOME OF INPATIENT PSYCHOTHERAPY FOR ADOLESCENTS (IPA) WITH PERSONALITY PATHOLOGY

Dine J. Feenstra, PhD, Elisabeth M. P. Laurenssen, MSc, Reinier Timman, PhD, Roel Verheul, PhD, Jan J. V. Busschbach, PhD, and Joost Hutsebaut, PhD

Little is known about the effectiveness of treatment programs for adolescents with personality disorders (PDs). This descriptive study investigated long-term outcomes of inpatient psychotherapy for adolescents (IPA). In addition, predictors of long-term treatment effects were investigated. Seventy adolescents who completed their treatment were followed during and after their stay in IPA. Semistructured interviews were used to assess Axis I and Axis II disorders. Patients completed questionnaires to measure symptom severity and personality styles and functioning. Patients showed improvements (d ranging from .18 to .80). After leaving IPA, a small relapse rate was shown. Higher levels of dependency or more Cluster C PD traits significantly predicted positive long-term treatment outcome. Although IPA might be effective for a particular group of adolescents, mean long-term treatment effects were at best modest. Given the high costs of IPA, developing adjustments to IPA and their guidelines is warranted to ensure cost-effectiveness.

According to DSM-IV-TR (American Psychiatric Association, 2000), personality disorders (PDs) can be diagnosed in adolescents. Recent research has shown that personality disorders are just as common in adolescents as they are in adults (see, for example, Feenstra, Busschbach, Verheul, & Hutsebaut, 2011; Grilo et al., 1998; J. G. Johnson et al., 2000; Westen, Shedler, Durett, Glass, & Martens, 2003), thus suggesting that adolescents are in great need of treatment as well, because they are at a particularly higher risk of developing a broad range of problems during adolescence. These problems include suicidal thoughts and attempts (Braun-Scharm,
1996; Westen et al., 2003), school problems (Westen et al., 2003), behavioral problems (J. G. Johnson et al., 2005), substance abuse (Serman, Johnson, Geller, Kanost, & Zacharapoulou, 2002), deviant sexual behavior (Lavan & Johnson, 2002), and emergency admissions (Kasen et al., 2007). They are also more likely to develop into vulnerable adults. Compared to adolescents without a disorder or to adolescents with only Axis I pathology, adolescents with personality disorders have an increased risk of failing at school (J. G. Johnson et al., 2005), mood disorders (Daley et al., 1999), anxiety and substance use disorders (J. G. Johnson et al., 1999; Levy et al., 1999), financial and health problems (Chen, Cohen, Kasen, & Johnson, 2006), more familial conflicts (J. G. Johnson, Chen, & Cohen, 2004), and interpersonal difficulties and stress (Daley, Rizzo, & Gunderson, 2006). Moreover, these adolescents report a low quality of life and high medical costs (Feenstra et al., 2012), as is true for adults with personality disorders (Soeteman, Hakkaart-van Roijen, Verheul, & Busschbach, 2008; Soeteman, Verheul, & Busschbach, 2008). Despite this compelling body of evidence, diagnosing personality disorders in adolescents remains controversial (Chanen & McCutcheon, 2008), and reports on effective treatments for this particular group of adolescent patients are scarce. Even less is known about the long-standing effects of these treatments and about which patient characteristics might predict long-standing treatment effects.

Chanen and colleagues (2008) were the first to describe long-term treatment effects for a group of adolescents with borderline personality disorder symptoms. They compared the treatment effects of cognitive analytic therapy (CAT) with good clinical care. No significant difference was found between the groups; however, some evidence suggested that the CAT group improved more rapidly. Schuppert et al. (2009, 2012) tested the effectiveness of an emotion regulation training (ERT) specifically developed for adolescents with symptoms of borderline personality disorder (BPD). Subjects were assigned to ERT plus treatment as usual (TAU) or TAU alone. In both treatment conditions, borderline symptoms reduced equally, while patients in the ERT plus TAU group also showed a significant increase in internal locus of control. Just recently, Rossouw and Fonagy (2012) presented the results of the first randomized controlled trial (RCT) investigating mentalization-based treatment (MBT) in adolescents who self-harm (70% of whom had BPD). MBT was compared to TAU and appeared more effective in reducing self-harm and depression. Other adaptations of MBT for adolescents have been described. Bleiberg (2001), for example, presented an adaptation of MBT based on developmental and attachment theory. Asen and Bevington (2007) have developed adolescent mentalization-based integrative therapy (AMBIT), which focuses on more “hard-to-reach” adolescents. Thus far, however, none of these treatments have been systematically evaluated. Backer, Miller, and van den Bosch (2009) identified seven studies investigating the effectiveness of dialectical behavior therapy (DBT) in adolescents. None of these were RCTs. Two of
these studies confirmed the effectiveness of DBT in adolescents with symptoms of BPD. Rathus and Miller (2002) investigated the effectiveness of DBT by comparing DBT to TAU and found larger effects in the DBT group. Fleischhaker, Munz, Böhme, Sixt, and Schulz (2006) studied the effectiveness of DBT in a pre-post test design and found that symptoms decreased during treatment. No long-term effects of these treatments have been reported.

All of the mentioned studies investigated outpatient treatment programs. Inpatient treatment has long been shown to be an effective treatment for adolescents (Fineberg, Sowards, & Kettlewell, 1980), but no research so far has focused on its effectiveness in the treatment of personality disorders in adolescents. The current study is the first to investigate the long-term effects of inpatient psychotherapy for adolescents (IPA). There is empirical support for inpatient psychotherapy for adults with personality disorders (Bartak et al., 2010, 2011; Chiesa, Fonagy, Holmes, & Drahorad, 2004; Gabbard et al., 2000; Vermote et al., 2009), and it is mentioned in the Dutch Multidisciplinary Guideline for Personality Disorders (Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling, 2008). Inpatient treatment for adult patients with personality disorders seems to produce improvements that last after treatment (see, for example, Gabbard et al., 2000; Vermote et al., 2009).

Adolescence is a developmental phase that is characterized by normative biological, cognitive, social, and emotional changes, such as maintaining friendships and forming new relationships within the family (Slot, 1994). Because adolescence is a developmental phase full of changes and transformations, we were particularly interested if treatment outcomes would last after the adolescents left the inpatient treatment program. We also wanted to explore which patients benefited the most from IPA after leaving the inpatient unit. This is relevant not only from a patient perspective, but also from a societal perspective because inpatient treatments are expensive.

Based upon previous studies and literature, we chose to include two theoretical variables that might predict treatment outcome for this type of inpatient treatment. First, we expected patients high on dependency to benefit more from IPA and, therefore, to show better outcomes at follow-up than patients high on self-criticism. This would be in line with the argument of Blatt and Felsen (1993) that patients with higher levels of dependency respond more effectively to a treatment program in which there is much room for personal interaction with the therapist. Patients with higher levels of self-criticism, on the other hand, may have more difficulties profiting from the therapeutic relationship. The rules and procedures inherent in an inpatient setting might conflict with their striving for autonomy. Second, we expected type of PD to be another predictor. In adult samples, it has been shown recently that Cluster B and Cluster C patients benefit from inpatient treatment; results, however, were more convincing for Cluster C patients (Bartak et al., 2010). We wanted to investigate if
these findings would also apply to adolescents with Cluster C PDs and if they would hold at follow-up measurements.

In this descriptive study, we investigated long-term outcome of IPA. In addition, this study examined potential predictors of long-term treatment effects.

**METHOD**

**PARTICIPANTS**

From June 2006 through January 2009, a total of 109 adolescents were admitted to the IPA unit of the youth department of de Viersprong Institute for Studies on Personality Disorders, Halsteren, The Netherlands, and enrolled in this study. De Viersprong is a highly specialized mental health care institute offering outpatient, day hospital, and inpatient psychotherapy for adolescents and adults with severe and complex personality pathology. In general, patients are referred to de Viersprong from all over the country because of complex pathology that appears to be refractory to outpatient treatment (mainly aimed at improving Axis I pathology). All patients underwent a standard assessment as part of the intake procedure, including semistructured interviews to measure Axis I and Axis II disorders, and completed several questionnaires. Patients, both treatment completers and dropouts, were asked to complete questionnaires at the start of treatment and at 6, 12, and 24 months after the start of treatment. All patients agreed to participate in the study and signed informed consent after the purpose of the study was explained. This study is part of the long-term outcome and process study of Treatment Refractory Adolescents with Personality Disorders (TRAP). Inclusion criteria for this study and admission to the inpatient unit were the presence of severe, chronic, and multiple complaints, leading to clinically significant distress and impaired social and school functioning, for which previous outpatient treatment has not resulted in significant improvement of functioning. Exclusion criteria were chronic psychotic disorders (e.g., schizophrenia), organic cerebral impairment, and mental retardation. Dropout was defined as any premature termination of treatment not mutually negotiated and agreed upon by staff and patient (Baruch, Gerber, & Fearon, 1998; Hatchett & Park, 2003; Richmond, 1992). Completion of treatment was defined as mutually agreed discontinuation of treatment (E. Johnson, Mellor, & Brann, 2009), with a minimum stay of 6 weeks. Of the 109 enrolled adolescents, five did not complete the entire assessment as part of the intake procedure. Thirty-four patients dropped out of treatment prematurely and were left out of the analyses, leaving 70 participants for the study sample. No patients were excluded due to the exclusion criteria. The study was approved by the Ethical Commission of the Department of Psychology of the University of Amsterdam.
IPA is an intensive treatment program and is modified along the lines of a therapeutic community approach. It incorporates mainly group-dynamic and milieu therapeutic approaches, in which adolescents are observed in and confronted with their dysfunctional relational, emotional, and behavioral patterns. Through the constant interaction with other adolescents, they can become more aware of their dysfunctional patterns. The basic approach of IPA is helping the adolescents to identify and explore their relational patterns, associated affects, and defensive functions. The therapeutic community provides a safe environment in which the adolescents can be confronted with these patterns, explore and discuss them, and practice new and more adaptive behavior. Safety and quality of the therapeutic community are monitored by a milieu therapist. The main goals of IPA are to address the maladaptive components underlying personality pathology (e.g., to gain more ego strength and better relational capacities) in order to reduce symptom severity and gain improvements in daily functioning.

Adolescents were placed in a group of a maximum of 10 adolescents. Individual psychotherapy sessions were offered once a week, and group psychotherapy sessions were scheduled three times a week. Individual and group sessions were given by a licensed psychotherapist who was responsible for the treatment process of that particular group. The techniques were mainly psychodynamically oriented, aimed at encouraging the expression and discussion of troubling emotions and distressing thoughts. When indicated (e.g., in cases of certain Axis I disorders such as obsessive-compulsive disorder), techniques or protocols from cognitive-behavioral therapy were used as well. Furthermore, nonverbal group therapies, such as psychomotor therapy and creative therapy, were offered four times a week. Other therapies were community-based group therapies with psychosocially trained nurses offered three times a week. In these therapies, adolescents were able to practice new behaviors, and they could translate these behaviors into daily life, including school and peer groups. Every adolescent had a mentor in the community setting (a psychosocial nurse) with whom he or she had individual sessions once a week (or more frequently if necessary). Psychiatric consults were scheduled on an as-needed basis. Sessions with a social worker were planned for the individual adolescent (and the family) to discuss plans for resocialization toward the end of treatment. A family therapist (licensed family therapist with several years of experience) had sessions with the adolescents and their families once every 2 or 3 weeks.

MEASURES
Screening Instruments. Anxiety and mood disorders were diagnosed using the Anxiety Disorders Interview Schedule for DSM-IV Child Version—Child interview (ADIS-C; Silverman & Albano, 1996; translated by Siebe-
The ADIS-C is a semistructured interview designed to measure anxiety and other Axis I disorders in children and adolescents. No interrater reliability data were collected in this study. However, research shows that the ADIS-C is reliable across time and informants and in comparison with other forms of assessment. Also, interrater reliability appeared to be good in a sample of children and adolescents ages 7–16 ($\kappa = .92$) (Lyneham, Abbott, & Rapee, 2007). The ADIS-C was supplemented by Sections E, G, and H of the Structured Clinical Interview for DSM-IV Axis I disorders (SCID-I; First, Spitzer, Gibbon, & Williams, 1997; translated by van Groenestijn, Akkerhuis, Kupka, Schneider, & Nolen, 1999) to diagnose substance-related disorders, somatoform disorders, and eating disorders, respectively. The SCID-I appears to have good interrater reliability ($\kappa = .85$), especially when interviewers received training (Ventura, Liberman, Green, Shaner, & Mintz, 1998).

The Structured Clinical Interview for DSM-IV Axis II Personality disorders (SCID-II; First, Spitzer, Gibbon, Williams, & Benjamin, 1996; translated by Weertman, Arntz, & Kerkhofs, 1996) was used to diagnose Axis II personality disorders. Criteria were scored if they were pathological, pervasive, and persistent and if they were present for 1 year, according to the guideline of the DSM-IV-TR. Because the DSM-IV-TR does not allow for antisocial personality disorder to be diagnosed in adolescents under the age of 18, this section was left out of the interview for adolescents under 18. Personality disorder not otherwise specified (PDNOS) was scored if a depressive personality disorder or a passive-aggressive personality disorder was present, or when at least 10 personality disorder traits from various disorders were scored without crossing the cutoff point of any formal PD. No interrater reliability data were collected in this study. Previous research has shown (see, for example, Maffei et al., 1997; Weertman, Arntz, Dreessen, Van Velzen, & Vertommen, 2003) that the DSM-IV version of the SCID-II has a good interrater reliability and test-retest interrater reliability for the presence or absence of a personality disorder diagnosis in adults. Although the SCID II is primarily designed for measuring personality disorders in adults, previous studies that included adolescent samples have shown that the SCID-II is a useful instrument with an adolescent group (Tromp & Koot, 2010).

**Outcome Measures.** Symptom severity, as reported by the adolescent, was measured by the Dutch version of the Brief Symptom Inventory (BSI; Derogatis, 1975; translated by de Beurs, 2006). It consists of 53 items covering nine symptom dimensions: somatization, obsession-compulsion, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism; and three global indices of distress: Global Severity Index, Positive Symptom Distress Index, and Positive Symptom Total. Respondents rank each feeling item (e.g., “your feelings being easily hurt”) on a 5-point scale ranging from 0 (not at all) to 4 (extremely). Rankings characterize the intensity of distress during the past 7 days. The Dutch version of the BSI showed good internal consistency,
with alpha scores ranging from .71 to .96. Test-retest reliability appeared adequate, with $r$s ranging from .71 to .90 (de Beurs, 2006).

Personality functioning was measured by the Severity Indices of Personality Problems (SIPP-118; Verheul et al., 2008). The SIPP-118 is a dimensional self-report measure and aims to measure the core components of (mal)adaptive personality functioning. The SIPP-118 is based on a theoretical view of personality pathology as constituting changeable components that can be generalized to various categories and types of personality pathology and that are continuous with normal, adaptive personality functioning. Examples of the changeable components of personality are the adaptive capacities, which refers to the dynamic organization of personality that concerns the regulation of self and relationships with others, and comprise characteristics such as affect regulation, self and other representations, identity, coping strategies, and acquired skills. The SIPP-118 asks the respondents to think about the past 3 months and to answer the extent to which they agree with statements such as “I frequently say things I regret later” or “Whenever I feel something, I can almost always name that feeling.” The response categories range from 1 to 4 and are described as “fully disagree,” “partly disagree,” “partly agree,” or “fully agree.” The measure comprises 16 facets clustered into five higher-order domains. The higher-order domains are named Self-Control (the capacity to tolerate, use, and control one’s own emotions and impulses), Social Concordance (the ability to value someone’s identity, withhold aggressive impulses toward others, and to work together with others), Identity Integration (the ability to see oneself and one’s own life as stable, integrated, and purposive), Relational Capacities (the capacity to genuinely care about others as well as feeling cared by them, to be able to communicate personal experiences, and to hear and engage with the experiences of others, often—but not necessarily—in the context of a long-term, intimate relationship), and Responsibility (the capacity to set realistic goals and to achieve these goals in line with the expectations you have generated in others). High scores indicate better adaptive functioning. Test-retest reliability, as measured over an interval of 14–21 days, was good, with $r$s ranging from .87 to .95. The SIPP-118 was tested in an adolescent sample and showed adequate psychometric properties, with alpha scores ranging from .62 to .89. The factor structure of the SIPP-118 adolescent sample was found to be similar to the factor structure found in adult samples, and the SIPP-118 appeared to discriminate between various populations. More pathological scores were found in the clinical sample than in a normal sample, and scores of adolescents with PDs were more pathological than those of adolescents without PDs (Feenstra, Hutsebaut, Verheul, & Busschbach, 2011).

The Dutch short version of the Depressive Experience Questionnaire for Adolescents (DEQ-A; Luyten, Corveleyn, & Blatt, 1997) was used to measure two personality styles: Self-Criticism and Dependency. The Depressive Experiences Questionnaire (DEQ; Blatt, D’Afflitti, & Quinlan, 1976)
was originally developed for adults. Items of the DEQ were rephrased to make them more appropriate for adolescents. A factor analysis showed three factors (Dependency, Self-Criticism, and Efficacy) similar to those in the adult sample (Blatt, Schaffer, Bers, & Quinlan, 1992). Blatt and colleagues (1992) found adequate internal consistency for the factors of the DEQ-A, with alpha coefficients ranging from .79 to .90. Furthermore, 1-year test-retest correlations for the three subscales ranged from .52 to .63. Respondents are asked to what extent they agree with the items. Responses to the DEQ-A are given on a scale from 1 (strongly disagree) to 7 (strongly agree).

Timing of Measurements. The screening instruments and all outcome measures were administered as part of the intake procedure. Patients completed the BSI and the SIPP-118 at 6, 12, and 24 months after the start of treatment. Patients completed the DEQ-A at 12 and 24 months after the start of treatment.

STATISTICAL ANALYSES

Differences at baseline variables between treatment completers and dropouts were investigated using chi-square tests and one-way ANOVAs. Results on outcome measures were examined using multilevel modeling. We used multilevel modeling to deal with (a) the dependency of repeated measures on the same subject in time, and (b) longitudinal data with observations missing at certain time points. To estimate treatment effects at 6, 12, and 24 months after the start of treatment, random intercept and slope models were postulated with time as Level I and patient number as Level II. The deviance statistic (Singer & Willett, 2003) was used to assess whether the slope could be postulated as random. Time, logarithm of time, and time squared were used as independent variables in the fixed part of the saturated models. Nonsignificant ($p > .05$) effects were excluded from the model until a final parsimonious model was reached that did not significantly differ from the saturated model. Subsequently, within-group effect sizes Cohen’s $d$ (Cohen, 1988) were calculated to describe changes from baseline to 24 months after the start of treatment.

In order to investigate clinically significant change in level of symptom severity, we computed the percentage of patients who achieved reliable change, the percentage who moved from a dysfunctional range to a normative range, and the percentage who had both reliable change and moved into a normative range as measured by the BSI (Jacobson & Truax, 1991). Reliable change was calculated using the following formula: $RC = 1.96 \sqrt{2(SE)^2}$, with $SE = 0.22$, $RC = 0.62$. A cutoff point for movement into a normative range was computed using the following formula: $(SD_{normal} \times M_{clinical} + SD_{clinical} \times M_{normal})/(SD_{normal} + SD_{clinical})$, with $M = .42$ ($SD = .40$) and $M = 1.21$ ($SD = 0.71$) for the normal and clinical populations, respectively (de Beurs, 2006). Clinical deterioration was also computed, defined as patients whose score on the BSI increased by the reliable change index.
Prediction of treatment outcome was investigated using the described multilevel analyses. Traits per cluster (A, B, C, and NOS) as well as their interaction with time, logarithm of time, and time squared were added to the saturated models. In a subsequent analysis, DEQ-A dimensions (self-criticism and dependency) and the interaction between these dimensions as well as the interaction with time, logarithm of time, and time squared were added to the saturated model.

RESULTS

PARTICIPANTS

Of the 70 adolescents, 58 were female (82.9%). The mean age for the study sample was 16.73 (range 14–19; SD 1.33). Baseline variables of the study sample and the excluded participants are presented in Table 1. No significant differences between the study sample and the excluded participants were found for demographic variables or Axis I disorders. As for Axis II disorders, personality disorder not otherwise specified (PDNOS) was more frequently classified in the study sample. As for the participants in the study sample, 11 had no diagnosis (15.7%), 25 adolescents (35.7%) had

| TABLE 1. Baseline Variables |
|-----------------------------|-----------------------------|
| **Study Sample**            | **Excluded Participants**   |
| **n = 70**                  | **n = 39**                  |
| **Demographic variables**   |                             |
| Female                      | 58 (82.9)                   |
| Age (M, SD)                 | 16.7 (1.33)                 |
| **Referred by**             |                             |
| General practitioner        | 9 (12.9)                    |
| Psychiatrist or psychologist with private practice | 10 (14.3) |
| General hospital            | 5 (7.1)                     |
| Mental health care center/ youth welfare | 46 (65.7) |
| **Treatment history**       |                             |
| Age at onset of problems (M, SD) | 12.51 (2.63) |
| Age at first treatment (M, SD) | 13.30 (2.88) |
| **Treatment duration**      |                             |
| Number of days in IPA (M, SD) | 339 (42.82) |
| **Clinical characteristics**|                             |
| Axis I disorders            |                             |
| Mood disorder               | 28 (40.0)                   |
| Anxiety disorder            | 26 (37.1)                   |
| Eating disorder             | 10 (14.3)                   |
| Somatoform disorder         | 2 (2.9)                     |
| Substance use disorder      | 7 (10.0)                    |
| Other disorder              | 1 (1.4)                     |
| Any Axis I disorder         | 55 (78.6)                   |
| Axis II disorders           |                             |
| Avoidant personality disorder | 11 (15.7)                  |
| Obsessive-compulsive personality disorder | 3 (4.3) |
| Borderline personality disorder | 14 (20.0) |
| Personality disorder not otherwise specified | 10 (14.3) |
| Any PD                      | 34 (48.6)                   |

Note. Data are presented as n (%), unless otherwise specified. The sum of the number of patients in the different diagnostic groups is higher than the total number of patients because patients can have more than one (personality) disorder. PD = personality disorder.
only an Axis I disorder diagnosis, 4 (5.7%) had only an Axis II personality disorder diagnosis, and 30 (42.9%) had both an Axis I disorder and an Axis II personality disorder diagnosis. Thus, of the 70 adolescents, 36 (51.4%) had no diagnosis on Axis II. Of these 36 adolescents, 4 (11.1%) met none of the criteria of a personality disorder. The other 32 (88.9%) met one or more criteria of a personality disorder, and 17 adolescents (47.2%) even met between five and nine criteria of a personality disorder. Because we used a cutoff point of 10 criteria for PDNOS for our study sample, this last group of adolescents did not receive a formal PDNOS diagnosis. However, Verheul, Bartak, and Widiger (2007) provided some appealing evidence that when a cutoff point of five criteria for PDNOS is used, it adds a large group of patients with a personality disorder with a similar level of functional impairment as the patients diagnosed with PDNOS using 10 criteria as a cutoff point. In this study, using a cutoff of five criteria for PDNOS, 17 patients could be added to the group with a personality disorder, leaving 6 adolescents (8.6%) without any Axis I or Axis II disorder diagnosis. In that case, 13 adolescents (18.6%) would have only an Axis I disorder diagnosis, 9 adolescents (12.9%) would have only an Axis II personality disorder diagnosis, and 42 adolescents (60%) would have both an Axis I disorder and an Axis II personality disorder diagnosis.

In both groups, most adolescents were referred by a mental health care center or youth welfare. Because these settings are so-called secondary care mental health settings (with general practitioner as a primary care setting), this gives some evidence for the severity of the pathology in these adolescents. All but one adolescent received treatment prior to their admission to IPA. As one can see in Table 1, adolescents sought help several years prior to their admission to IPA.

TREATMENT OUTCOME

The adolescents showed significant improvement on all outcome measures. They reported significantly less symptom severity and significantly better personality functioning 24 months after the start of treatment. The estimated outcome results are presented in Table 2. Effect sizes at 24 months after the start of treatment range from .18 to .80, indicating small to large effects.

CLINICALLY SIGNIFICANT CHANGE

As measured by the BSI, 45.0% of the adolescents (n = 18 of 40) had reliable change. Thirty-five percent of the adolescents (n = 14) moved from a dysfunctional range to a normative range of symptom severity, whereas 27.5% demonstrated both reliable change and movement into the normative range (n = 11). Four adolescents (10.0%) showed clinical deterioration.
PREDICTION OF TREATMENT OUTCOME

Predictive Value of Personality Disorder Traits per Cluster. Table 3 shows the predictive value of personality disorder traits per cluster on level of symptom severity. Adolescent patients with more Cluster C PD traits started with higher levels of symptom severity and ended with lower levels of symptom severity, as opposed to adolescent patients with fewer Cluster C PD traits (see Figure 1).

Predictive Value of Dependency and Self-Criticism. In Table 4, the predictive value of levels of dependency and self-criticism on level of symptom severity is shown. Adolescents with higher levels of self-criticism have higher levels of symptom severity at all time points. Furthermore, patients with higher levels of dependency improve more rapidly in terms of symptom severity (see Figure 2).

DISCUSSION

In this descriptive study, we investigated long-term treatment effects in adolescents with personality pathology admitted to IPA. Our results indi-
cate that these adolescents show improvement in both level of symptom severity and personality functioning 24 months after the start of treatment. Effect sizes (Cohen’s $d$) at follow-up ranged from .18 for Relational Capacities to .80 for Self-Control. Patients who benefited most from IPA were characterized by higher levels of dependency or more Cluster C PD traits. Almost 30% of the adolescents showed clinically significant change as well as movement into a normative range 24 months after the start of treatment.

The results for the group of adolescents as a whole were slightly disappointing. While a minority of adolescents made significant progress in terms of symptom severity, overall progress was small because a large group of adolescents did not change or showed only little improvement in level of symptom severity. The group showed better levels in most domains of personality functioning and lower levels of symptom severity at both 12 and 24 months after the start of treatment. However, the adolescents were not able to maintain the gains they had made at the end of treatment. A small relapse in results was visible for most outcome measures. This is in contrast to the findings of, for example, Chanen and colleagues (2008), who found ongoing improvements after the end of treatment. Sim-

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>$SE$</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.374</td>
<td>.310</td>
<td>1.207</td>
<td>.230</td>
</tr>
<tr>
<td>time linear</td>
<td>−.028</td>
<td>.020</td>
<td>−1.410</td>
<td>.161</td>
</tr>
<tr>
<td>time squared</td>
<td>.002</td>
<td>.000</td>
<td>3.343</td>
<td>.001</td>
</tr>
<tr>
<td>Dependency</td>
<td>−.047</td>
<td>.060</td>
<td>−.778</td>
<td>.438</td>
</tr>
<tr>
<td>Self-criticism</td>
<td>.302</td>
<td>.054</td>
<td>5.552</td>
<td>.000</td>
</tr>
<tr>
<td>time linear × Dependency</td>
<td>−.006</td>
<td>.003</td>
<td>−2.056</td>
<td>.041</td>
</tr>
</tbody>
</table>

Note. $SE$ = standard error. The slope ($t$) was not postulated as random in this model.
ilar findings were also reported by Svartberg, Stiles, and Seltzer (2004), who found ongoing treatment effects for short-term outpatient treatment in adult patients with personality pathology.

Several factors might explain our modest results. First, IPA was a non-manualized treatment program that was developed mainly for a less disturbed patient group. Its focus was on the supposed mechanisms of change in what one would call a "neurotic" patient group. However, it lacked a specific theoretical frame to address the personality pathology of adolescents with more severe PDs. As such, it was not specifically developed for treating less integrated and more disturbed PD adolescents. The focus of the team was less on providing consistency and continuity throughout the program, and it assumed that patients themselves were able to integrate the different components of treatment. This might make a therapeutic community–based program like IPA less suited for more severely disturbed adolescents. As Bateman and Fonagy (2004) stated, for such patients "it is crucial to maintain consistency, constancy, and coherence of treatment because individuals with PDs detect and exploit inconsistency" (p. 187).

This was not explicitly the focus in IPA, probably causing more inconsistencies in a multidisciplinary team in the absence of a manual than is appropriate for these PD patients. Remarkable differences were found in effect sizes for the different outcome variables. This might suggest that IPA serves certain goals better than others. One might speculate that the structured nature of IPA and the constant confrontation might have led to better Self-Control and Social Concordance, as was presented by higher effect sizes on these domains of the SIPP-118. On the other hand, the lack of a consistent theoretical treatment model aimed at targeting the under-
lying mechanisms of personality pathology might explain, for example, the minimal changes in more structural changes in relational functioning.

Second, the inpatient setting might have given the adolescents a temporary escape from the difficulties they were experiencing with developmental tasks (e.g., school, peers, and parents). An inpatient setting like IPA provides a level of containment and may reverse a downward developmental spiral by providing a refuge that may be beneficial in the short term. It could be that the changes in the adolescents’ levels of personality functioning were not anchored enough to cope with the difficult tasks still waiting for them at home.

However, some nuance is provided by the patient characteristics we investigated. We found that higher levels of dependency significantly predicted treatment effects at follow-up. Adolescents who scored higher on the dependency dimension at the start of treatment improved more rapidly in terms of symptom severity and even showed ongoing improvement after treatment. This is in line with previous research showing that dependent patients benefit more from treatment programs in which there is much room for personal interaction with the therapist (Blatt, 2008; Blatt & Felsen, 1993; Blatt & Luyten, 2009). Research has shown that the two dimensions of Blatt’s model, self-criticism and dependency, are closely linked to the two dimensions underlying attachment style, avoidance and anxiety, respectively (Luyten & Blatt, 2011, 2012). Attachment avoidance, related to self-criticism, is defined as “discomfort with closeness and with discomfort depending on others.” Attachment anxiety, related to dependency, is defined as “fear of rejection and abandonment” (Mikulincer & Shaver, 2007). When stress increases, patients high on self-criticism will more likely rely on themselves, whereas dependent patients will rely on others. The supportive nature of IPA, in which there is much room for interpersonal contact, might therefore be more adequate for dependent patients. These patients might also be the ones who can rely on others for support after their treatment, which might be an explanation for the ongoing improvements in level of symptom severity.

We found that adolescent patients with more Cluster C PD traits showed more improvements in level of symptom severity. This is in line with the research of Bartak and colleagues (2010), who found beneficial outcomes for adult Cluster C patients in an inpatient setting. On the one hand, it could be expected that Cluster C patients profit more from treatment than patients with personality disorders from other clusters. Kernberg and Calidor (2005), for example, suggest that most Cluster C personality disorders are situated at the higher end of the continuum used to describe levels of personality organization. However, even when Cluster C patients do not function at the neurotic level of personality organization, they are likely to have less splitting and less severe polarization between positive and negative, leading to better interpersonal functioning and less identity diffusion than may be the case with Clusters A and B. They therefore might benefit more than other (Cluster A or B) patients from the treatment
offered with IPA. Furthermore, Milrod and colleagues (2007) found a moderating effect of Cluster C personality disorders on the outcomes of a psychodynamically oriented psychotherapy focused on panic disorders. Patients with a Cluster C personality disorder had more beneficial outcomes than patients without Cluster C personality disorders. On the other hand, other studies suggest a poorer outcome for Cluster C patients (see, for example, Oleski, Cox, Robinson, & Grant, 2012). It might be worthwhile to investigate whether other patient characteristics might moderate the outcomes of Cluster C patients.

Our findings elicit some important issues. Given the expensive nature of a treatment like IPA and the slightly disappointing results, one could consider whether patients with personality disorders are better treated in an outpatient, and presumably less expensive, setting. An outpatient, step-up or step-down trajectory of IPA might provide more opportunities for the adolescents to practice in their own environment and with their families what they have learned in treatment. Moreover, IPA was a nonmanualized treatment program. In the literature on adult personality pathology, several manualized treatment programs have proven to be effective for patients with personality disorders (e.g., mentalization-based treatment (MBT; Bateman & Fonagy, 2004), transference focused psychotherapy (TFP; Clarkin, Yeomans, & Kernberg, 2006), dialectical behavior therapy (DBT; Linehan, 1993), or schema focused therapy (SFT; Young, Klosko, & Weishaar, 2003). It might be worthwhile to adapt such models to adolescents in order to see whether these models might also lead to benefits for younger patients. The results concerning patient characteristics, however, suggest that not all adolescents might profit from the same interventions in the same way. Blatt (2008) states that both dependent and self-critical patients struggle with distorted and pathological introjects. These introjects have to be addressed in therapy. In this way, therapy for both dependent and self-critical patients follows the same line. The way in which these introjects are addressed might, however, be different for both patient groups. It seems important that therapists are aware of the different trajectories these patients will have to follow in order to gain maximum treatment effects. Although our study was not conducted with this aim in mind, the results nevertheless shed an interesting light on the actual discussion of severity and traits in predicting outcome. Our results suggest that severity of (personality) pathology does not predict treatment outcome. If it would, one would expect patients with a substantial number of personality disorder traits, as well as patients high on both dependency and self-criticism, to show less beneficial outcome. In this study, severity was measured by the total number of traits and not by the nature of the traits. We cannot rule out that using another definition of severity would yield different results. Our results suggest, however, that patient characteristics are a valuable factor in predicting treatment outcome.

This study has several strengths and limitations. A strength is the performance of thorough assessments of Axis I and Axis II disorders. A limi-
tation is that we did not include a control group in this study, which complicated the interpretation of the observed long-term treatment effects. We cannot rule out that no treatment or, in other words, the natural course of the pathology would yield similar or even larger improvements compared to IPA. Furthermore, we had only one treatment modality and could not differentiate between short- and long-term inpatient programs as Bartak and colleagues (2010, 2011) did. In that respect, further research is needed to evaluate different treatment modalities. Additionally, in our sample, females were overrepresented, and more research is needed for adolescent males. Furthermore, our sample consisted mainly of native Dutch adolescents with minimal diversity in cultural background. This might limit the generalizability of our findings to adolescents with other cultural backgrounds. Further research is needed that includes adolescents from other cultural contexts. Despite substantial attempts to involve all adolescents in this study, a large group of adolescents did not complete the questionnaires at all time points. To account for these missing values, we used multilevel modeling. Further research should employ even more effort to ensure compliance. Finally, in this study we used questionnaires initially designed for adults. There are now questionnaires specifically developed for assessing personality pathology in adolescents. Future studies should use these resources.

Our results show that adolescents with personality disorders admitted to an inpatient treatment program show less symptom severity and better personality functioning 24 months after the start of treatment. Our study shows that especially patients with higher levels of dependency or patients with more Cluster C PD traits profit most from this particular inpatient treatment program. Overall, long-term effects found in this study were slightly disappointing because inpatient treatment is expensive. Further research is needed to investigate for which patients an inpatient treatment program may be more (cost-)effective than outpatient treatments in order to shed light upon the position of inpatient treatment in the overall system of care for adolescents with personality disorders.

REFERENCES


