Beyond detention: The effectiveness of aftercare for juvenile and young adult offenders

James, C.

Citation for published version (APA):
Chapter 5.

The effectiveness of the New Perspectives Aftercare Program for juvenile and young adult offenders: Recidivism outcomes.

This chapter is submitted for publication as:
Abstract

Aftercare can be viewed as the pinnacle of the reentry process, yet little is known about its working mechanisms and what works for whom under which circumstances. This study examined the effectiveness of the New Perspectives Aftercare Program (NPAP) for serious juvenile and young adult offenders in The Netherlands. The aim was to determine whether NPAP was effective in reducing recidivism and what factors moderated the overall effectiveness. Participants ($N = 161$) were randomly assigned to NPAP ($n = 79$) or existing aftercare services (‘treatment as usual’, $n = 82$). Recidivism data were based on official delinquency records. Survival analysis was used, with an average follow-up period of respectively $M = 2.42$ years ($SD = .75$) for the NPAP group and $M = 2.33$ years ($SD = .74$) for the TAU group. The two conditions were compared in terms of time to re-conviction, and frequency of re-arrests and violent or non-violent rearrests, using chi square-tests and univariate ANOVA tests. No evidence was found to support the claim that NPAP was more effective in reducing recidivism than treatment as usual. Results showed that age of onset and number of prior offenses increased the risk for recidivism. Although the current study contributes to the small, but important body of empirical work, much remains to be learned on how to successfully reduce recidivism among this group of serious young offenders. Further research is needed to identify working mechanisms and possibilities to improve and strengthen aftercare programs, such as NPAP.
The majority of the juvenile and young adult offenders who reenter society every year after having spent time in secure care or (juvenile) detention do not manage to break the cycle of crime, and return into the criminal justice system within the first few months and years post-release (Abrams & Snyder, 2010; Spencer & Jones-Walker, 2004; Wartna et al., 2012). It appears to be difficult for them to turn their lives around and mature into being responsible and law-abiding adults (Snyder & Sickmund, 2006).

Some scholars argue that correctional rehabilitation programs alone are not enough to reduce the risk for recidivism, and claim that they should be accompanied by adequate reentry and aftercare services (Abrams, 2006; Steinberg, Chung, & Little, 2004). Research has shown that lower recidivism rates and positive adjustment to the community are achieved when the transition from correctional facilities to the community is directed and supervised (Fagan, 1990; Goodstein & Sontheimer, 1997). Therefore, over the past decades, various aftercare programs have been developed for juvenile and young adult offenders in an effort to improve successful re-entry and reintegration and prevent further offending. Yet, not many of them have been accompanied by thorough scientific research into their effectiveness, as was shown in a recent meta-analytic review by James, Stams, Asscher, De Roo, and Van der Laan (2013). Their study revealed that the overall effect size for aftercare programs was small, indicating that aftercare programs generally had a modest effect on recidivism reduction. Aftercare appears to be most effective if it is well-implemented, intensive and consists of individual instead of group treatment, aimed at older and high-risk youth.

One evidence-based aftercare program, which has been developed during the past few years, is the New Perspectives Aftercare Program (NPAP), an intensive, multimodal, community-based and offender-focused intervention in The Netherlands, starting reentry services three months pre-release from a secured institution and continuing aftercare until 6 months post-release. NPAP seems to be a promising aftercare intervention, since it is aimed at high risk youths, targets specific criminogenic needs, such as criminal thinking patterns, cognitive distortions and prosocial (coping) skills, which have been shown to have a significant impact on delinquency (Andrews & Bonta, 2010; Lösel & Beelmann, 2003; Nas, Brugman, & Koops, 2008; Rokach, 1997, 2000; Philips & Lindsay, 2011).

The main goal of the current study was to examine whether NPAP produces outcomes that are superior to existing services (‘treatment as usual’) by means of a randomized controlled trial. The primary outcome was a decrease in the frequency and seriousness of criminal behavior following referral to aftercare. A second aim was to examine moderators of the effectiveness of NPAP because research on the effectiveness of interventions is moving away from merely focusing on whether or not treatment works towards more clinically relevant ques-
tions: for whom does the intervention work and under what conditions (Kraemer, Wilson, Fairburn, & Agras, 2002)?

The number of prior offenses and the distinction between the number of violent and non-violent prior offenses were included as potential moderators of program effectiveness because history of criminal offending is one of the most important (static) risk factors of reoffending in delinquent youth (Andrews & Bonta, 2010; Grieger & Hosser, 2014). Another important static risk factor of recidivism is age of first arrest (Loeber, & Farrington, 1998), which was also included as a moderator.

Furthermore, ethnic minority status might influence program effectiveness. There is empirical evidence showing that ethnic minority groups are over-represented in the population of juvenile delinquents and youth care globally (Blom & Van der Laan, 2006; Junger, 1990; Snyder et al., 2006). Whether clients of ethnic minority groups should be treated with the same methods and interventions as the majority population of a particular country has been much discussed and studies have shown equivocal results (de Anda, 1997; Van der Put, Stams, Deković, Hoeve, & Van der Laan, 2013; Wilson, Lipsey, Soydan, 2003). By including ethnic minority status as a moderator, we can establish whether the aftercare program can be offered to both groups regardless of their ethnic background.

In addition, age of the offenders was included as a moderator, since prior research on aftercare programs has shown that interventions are more effective when aimed at older youths (James et al., 2013), likely because of their readiness for change connected to their moral, social and identity development, which continues into young adulthood (Arnett, 2000). Furthermore, some researchers have argued that little benefit can be derived when individuals are forced into treatment by the criminal justice system (Hartjen, Mitchell, & Washburne, 1981; Parhar, Wormith, Derkzen, & Beauregard, 2008; Platt, Buhringer, Kaplan, Brown, & Taube 1988). According to this view, it is a poor investment to devote resources to youths who are unlikely to change because they are only externally motivated to change by the possible negative consequences of non-compliance. Based on this notion, whether or not youths were referred to the aftercare program on a voluntary or a mandatory basis was included as a potential moderator of program effectiveness.

Finally, treatment retention was included as a potential moderator, since prior research has shown that whether or not participants dropped out of treatment is a good predictor of an individual’s long-term success (Hubbard, Craddock, Flynn, Anderson, & Etheridge, 1997; Simpson, Joe, & Brown, 1997).

In sum, not many studies aimed at finding effects of aftercare programs on delinquency and even less aftercare programs have been evaluated rigorously. Of the studies including criminal recidivism as an outcome, relatively few (e.g.,
Bergseth & McDonald, 2007; Gray et al., 2005; Wiebush, Wagner, McNulty, Wang, & Le, 2005) have examined moderators of effectiveness. Therefore, the present study focuses on these understudied topics in research on aftercare programs for juvenile and young adult offenders.

Method

Participants

Participants were juveniles referred to NPAP between January 2010 and July 2012 and who met the inclusion criteria of NPAP: 1) a minimum of four weeks detention/secure care; 2) three criminal acts, including one serious offense; 3) moderate to serious problems regarding social skills, criminal thinking patterns and behavior and 4) a medium to high recidivism risk (Vogelvang & Schut, 2008). A sample size of 64 per treatment condition is sufficient to test the hypotheses assuming .80 power, an alpha of .05, and a medium effect size (Cohen, 1992). However, taking into account possible attrition, inclusion continued until a minimum sample size of 80 participants per treatment condition at baseline was reached.

A total of $N = 163$ were recruited for the study at baseline and randomly assigned to the intervention NPAP group ($n = 79$) or the control group ($n = 84$). Two participants (both TAU) had to be excluded from the analyses of the official judicial data, because one participant could not be found in the Research and Policy Database for Judicial Documentation that was used and the other had no history of criminal offending, resulting in $N = 161$, with $n = 79$ for the NPAP and $n = 82$ for the TAU group. Figure 1 depicts the flow of participants for the complete study.

The majority of the sample was male ($n = 156$). The sample ranged in age from 15 to 26 years old ($M = 19.69$, $SD = 2.47$) at index date (start of recidivism measurement). In this study, 86.7 % had an ethnic minority status, meaning that at least one of the youth’s parents was born outside of The Netherlands. The largest ethnic minority group (44.3 %) consisted of second generation non-western immigrant workers (from Moroccan and Turkish descent). According to official judicial records, participants were on average 15.8 years old ($SD = 5.04$) when they first came into contact with the police and had a criminal record history of on average 5.47 offenses ($SD = 3.75$) prior to referral to NPAP. An overview of the key background variables of the experimental and control group is shown in Table 1.
Independent samples $t$-tests for continuous variables and chi-square analyses for categorical variables were conducted to examine differences between treatment conditions on demographic and criminal history variables. No significant differences were found on any of these variables, suggesting that randomization was successful.

**Procedure**

The study involved random assignment of individuals referred to NPAP, equally divided into an experimental (NPAP) condition and control condition (‘treatment as usual’, TAU).

Individuals were recruited after being referred to NPAP by one of various possible referral agencies, such as a youth care institution or probation officer. The study included three major cities in The Netherlands: Amsterdam, Utrecht and The Hague. After a client was referred to NPAP, he/she was screened by the NPAP’s team manager and/or behavioral scientist based on an intake form and the available case file. Juveniles and young adults aged 16 to 24 were eligible for the aftercare program and inclusion in the study if they 1) had spent a minimum of four weeks in detention; 2) had committed a minimum of three criminal acts, including one serious offense; 3) moderate to serious problems regarding social skills, criminal thinking patterns and behavior and 4) had a medium to high recidivism risk, based on a recent risk assessment by the RISc$^{13}$ (Adviesbureau Van Montfoort & Reclassering Nederland, 2004) or the SAVRY$^{14}$ (Lodewijks, Doreleijers, Ruiter & Wit de - Grouls, 2006). Exclusion criteria were a low IQ (<75); severe psychiatric problems and/or; substance abuse problems; and lack of motivation (Vogelvang & Schut, 2008). If a client was found eligible for enrolment in NPAP, he or she was informed about the trial by the referral agency. When informed consent was obtained, computerized randomization took place.$^{15}$ The medical ethical committee of the University of Amsterdam approved the design of the study (number 10.17.1193).

To establish whether the participants had reoffended since being referred to NPAP official records were requested from the Judicial Information Service [Justitiële Informatiedienst, JustID]. The data were provided from the Research and

---

$^{13}$ RISc stands for ‘Risico Inschattingsschaal’ [Risk Assessment Tool] and is the Dutch adaptation of the Offender Assessment System (OASys) (OASys Development Team, 2001).

$^{14}$ The SAVRY is the Dutch translation of the Structured Assessment for Violence Risk in Youth (Borum, Bartel, & Forth, 2002).

$^{15}$ The complete research protocol can be obtained from the first author.
Policy Database for Judicial Documentation [Onderzoek- en Beleidsdatabase Justitiële Documentatie, OBJD], held by the Research and Documentation Center of the Netherlands Ministry of Security and Justice [Wetenschappelijk Onderzoek- en Documentatiecentrum, WODC], that coded the data according to their Recidivism Coding System (RCS) (Wartna, Blom, & Tollenaar, 2011). Data on delinquency were based on official records, both juvenile and adult records, up to July 2013.

Since we did not have exact data on the detention period preceding aftercare or the release date of the offender, the starting point of the observation period – the period for which the recidivism is set – is the date on which a person commenced with NPAP or TAU. For the non-starters, we took the date of referral, added by the mean (in days) between referral and start date for the juveniles who did enroll in NPAP or TAU. Recidivism is the date on which the index case is registered with the Public Prosecutions Service (OM). That way we depict recidivism from the moment the interference of the criminal justice system starts. All cases that were registered before the index date were classed as previous convictions.

Formal consent for the recidivism study was obtained from the Netherlands Ministry of Security and Justice.

**Conditions**

**New Perspectives Aftercare Program (NPAP).** The New Perspectives Aftercare Program (NPAP) is an intensive aftercare program for serious juvenile and young adult offenders, aged 16 to 24 years, reentering society after a period of detention or secure care, with a moderate to high recidivism risk. The goals of the treatment are the prevention of recidivism (new convictions) by modifying cognitive distortions and behavior and improving social skills. The intervention is intensive, vigorous and highly individualized; combining reintegration in the neighborhood, work or school, and focusing on the network of the juvenile or young adult as a whole. The intervention strategies include coordinated case management, motivational interviewing and cognitive behavioral interventions focused on controlling impulses, problem solving and criminogenic thinking patterns (Vogelvang & Schut, 2008).

NPAP is divided into three phases and it is typically delivered for 9 months, with a 3-month follow-up period. The first (reentry) phase commences during the last three months of detention to ensure a smooth transition from the closed environment back into the community, followed by an intensive phase of three months and a consolidation phase. During the intensive phase, the youth care workers have 6-8 hours a week per client. In addition, they have low caseloads.
(six to seven clients maximum), are available 24 hours a day, are outreaching, invest in a therapeutic relationship with their clients, and aim to build a prosocial network that the juveniles and young adults can rely on beyond the intervention period.

**Treatment as Usual (TAU).** The participants assigned to the control group were offered the usual services for this target group, a broad array of social and mental health interventions, including juvenile and adult justice services, child welfare services, and youth care services. Predominantly, these services included individual treatment, with 68.3 % of the TAU group youths receiving individual counseling and case management through a secondary prevention program, and 18.3 % through a community-based intervention. Some youths were referred to family-based interventions (FFT or MDFT, 3.7%) and treatment in a residential setting (3.7 %). Most of the interventions provided as treatment as usual consisted of components that are also present in NPAP and focused on similar risk factors. Of the control group youths, 32.9 % eventually received no treatment, mostly due to a lack of motivation, versus 21.5 % of the NPAP group.

**Measures**

The demographic data and details of the treatment were obtained through file analysis and self-report. The following demographic variables were coded: gender, age (at index date), and ethnicity. In addition, it was coded whether or not juveniles commenced and completed NPAP or TAU and if they were referred to aftercare on a voluntary or mandatory basis. The following variables regarding criminal history were coded: age at first offense, total number of prior offenses, total number of prior violent offenses, and total number of prior non-violent offenses (see Table 1).

As described in the procedure, our measures of recidivism were based on the Recidivism Coding System (RCS) (Wartna et al., 2011). Recidivism was defined as the occurrence of any new conviction for any criminal offense after the start of NPAP or TAU. Recidivism was measured in four different ways to in-

---

16. Considered as new convictions were cases disposed of by the Public Prosecutor’s Service (excluding dismissal by reason of unlikelihood of conviction and cases that are transferred to another district), cases ending in a guilty verdict by the judge and cases that are not yet decided upon. The latter category was taken into account in determining the recidivism rate, based on the experience of the Research and Documentation Center of the Netherlands Ministry of Justice, that only a relatively small percentage of criminal cases end in acquittal, dismissal by reason of unlikelihood of conviction or a technical
clude important nuances in reoffending patterns. Firstly, recidivism was treated as a dichotomous variable (whether or not convicted for any new offense; any violent offense and/or any non-violent offense at least once). In addition, frequency of reoffending was taken into account as a continuous measure (number of any reconvictions; any violent reconviction and/or any non-violent reconviction). Furthermore, we included velocity of recidivism (time until first re-conviction; time until any violent reconviction and time until any non-violent reconviction). In accordance with the RCS guidelines, minor offenses, like traffic offenses, were not taken into consideration. Cases ending in acquittal, dismissal by reason of unlikelihood of conviction or clearance of charges by the court were also not taken into account.

**Analytic strategy**

We examined official delinquency outcomes up to July 2013. We compared the two conditions in terms of time to re-conviction, and frequency of re-arrests and violent or non-violent rearrests, using chi square-tests and univariate ANOVA tests. Survival analysis (based on the Kaplan-Meier procedure; IBM SPSS Version 20.0 for Windows) was used to obtain the cumulative survival functions (or survival curves) for participants who were randomly assigned to the NPAP or TAU group, whose average follow-up periods were respectively $M = 2.42$ years ($SD = .75$) and $M = 2.33$ years ($SD = .74$), and did not significantly differ between conditions, with $t (1, 159) = .78, p = .436$. The cumulative survival function represents the proportion of participants who survived any type of reconviction; any violent and/or any non-violent reconviction (i.e., were not reconvicted) in each group by the length of time (in days) from moment of commencing aftercare.

We examined the whole follow-up period for the total sample with Cox Regression analyses to examine differences in the prevalence of total recidivism; violent and non-violent recidivism between the NPAP and TAU group. To determine an effect size for the survival function, we performed a Cox proportional hazards regression (Cox, 1972).

For the moderator analyses on the recidivism data, the same Cox regression analyses were conducted, in which potential moderators were entered in the first step, to establish which covariates were significantly related to the survival time, decision by the court. Nine out of ten cases brought to the attention of the Public Prosecution Service end in a ‘valid’ disposal (Wartna, Blom & Tollenaar, 2011).

17. Violent offenses also included sexual offenses and property crimes with the use of violence. Non-violent offenses included property crimes, public order offenses, drug offenses, traffic offenses and other crimes.
and condition was added in the second step to establish whether condition predicted survival time after statistically adjusting for the covariates in the equation (Tabachnick & Fidell, 2013). In addition, to determine whether there are certain types of offenders for whom aftercare may be more or less effective, moderator *condition interaction terms were conducted by adding the interactions in the third step.

Results

Number of post treatment reconvictions and time to first recidivism

In order to examine differences in number of reconvictions and time to recidivism, Chi Square tests and ANOVA’s were conducted (Table 2). The findings indicate that there were no differences between the NPAP and TAU in frequency of reconvictions at follow-up, whether considered overall, or for violent or non-violent offenses separately. Neither were there any differences in time to any recidivism, nor for time to first violent and non-violent reconviction.

Survival functions

Figure 2 shows the survival curve of the juveniles that reoffended across the total follow-up period separately for the experimental and control group. The horizontal axis represents time in days until recidivism, and the vertical axis represents the cumulative proportion of those at risk for recidivism for those who have not recidivated. A log-rank test (with the Kaplan-Meier estimator; Kaplan & Meier, 1958) revealed that the survival functions for the two groups were not significantly different, $\chi^2(1, N = 161) = 1.601, p = .206$. Cox regression analysis showed that at the end of the follow-up period (on average 2.37 years), 79.7 % of the NPAP group and 69.5 % of the TAU group had recidivated at least once. As was shown by the log-rank test, Cox proportional hazards regression analysis confirmed that the differences between the NPAP group and the youths who received treatment as usual regarding the level of recidivism were non-significant, with a Hazard Ratio for condition of .794, $p = .207$, 95 % CI [.554 to 1.136].

Another set of survival analyses was conducted to examine between-groups differences on time to first arrest for various types of offenses. Of the participants who recidivated ($N = 120$), respectively 65.1 % of the NPAP group ($n = 63$) and 59.6 % of the TAU group ($n = 57$) committed one or more violent crimes, and 71.4 % of the NPAP group and 78.6 % of the TAU group committed one or more non-violent crimes. As depicted in Figures 3 and 4, respectively, participants in
the NPAP group were not significantly lower at risk for reconviction for violent offenses, $\chi^2 (1, N = 120) = .005, p = .943$; and nonviolent offenses, $\chi^2 (1, N = 119) = .668, p = .414$, during follow-up than were participants in the TAU group, with Cox proportional hazards ratios of respectively, $984, p = .943, 95 \% CI [.623 to 1.552]$ and $1.190, p = .414, 95 \% CI [.784 to 1.806]$.

**Potential moderators of criminal recidivism in NPAP**

To examine whether the effectiveness of NPAP was influenced by one or more factors, moderator analyses were conducted, with age, age at first arrest, ethnicity, voluntary vs. mandatory treatment, whether or not juveniles commenced and completed aftercare or TAU, and number of prior offenses as potential moderators. These were entered as covariates in the first step, with condition entered in the second step and the interaction between condition and the potential moderators in the third step.

The results of the Cox regression analysis showed that there was no statistically significant effect of treatment condition on overall recidivism, violent recidivism and non-violent recidivism after adjusting for all the covariates, with respectively $\chi^2 (1, N = 158) = 1.998, p = .157; \chi^2 (1, N = 158) = 1.471, p = .225$; and $\chi^2 (1, N = 158) = 1.445, p = .229$.

However, of the set of potential moderators, a history of criminal offending did increase the risk for general reoffending. The number of prior offenses significantly predicted survival time, Wald $(1, N = 158) = 12.988, p < .001$, with a Hazard Ratio of $1.137, 95 \% CI [1.060 to 1.220]$. This indicates that each increase of one prior offense increases the risk for reoffending by $13.7 \%$. Yet, no significant interaction effect was found for condition * number of prior offenses on overall recidivism, with a Hazard Ratio of $0.987, p = .856, 95 \% CI [.857 to 1.137]$.

Furthermore, results showed that the number of prior offenses increased the risk for non-violent reoffending, Wald $(1, N = 158) = 17.289, p < .001$, with a Hazard Ratio of $1.189, 95 \% CI [1.096 to 1.290]$, indicating that each increase of one prior offense increases the risk for non-violent reoffending by $18.9 \%$. Age at first arrest also significantly predicted survival time, Wald $(1, N = 158) = 4.497, p < .05$, with a Hazard Ratio of $1.168, 95 \% CI [1.012 to 1.347]$, indicating that the older participants were when they first came into contact with the criminal justice system, the more likely they were to reoffend non-violently. Yet, no significant interaction effects were found on non-violent recidivism for either of the two moderators, with a Hazard Ratio of $0.991, p = .917, 95 \% CI [.833 to 1.178]$ for condition * number of prior offenses and a Hazard Ratio of $0.959, p = .788, 95 \% CI [.704 to 1.305]$ for condition * age at first arrest.
Thus, survival time was predicted by several factors, but not by condition or by interaction between moderator and condition. Increases in risk for overall recidivism; violent recidivism and non-violent recidivism were, respectively, associated with the number of prior offenses committed and in case of non-violent reoffending by the number of prior offenses committed and age at first arrest. The other potential moderators: ethnicity, program retention and whether or not participants were referred to aftercare on a mandatory or voluntary basis did not have any effect on the recidivism outcomes.

Discussion

The present study examined the effectiveness of the aftercare program NPAP for juvenile and young adult offenders by analyzing the judicial data from participants who were randomly assigned to the intervention or control group that received treatment as usual.

The findings reported here suggest that NPAP failed to significantly affect overall recidivism and is not more effective in reducing recidivism during the period immediately following incarceration than treatment as usual. The results showed that there were no significant differences between the intervention and control group youths in their frequency and seriousness of reoffending as measured by post-release reconvictions for criminal offenses. The risk for recidivism was similar in both NPAP and TAU group youths. Both bivariate statistics and survival analysis revealed that both groups recidivated at an equivalent high rate and interval, with approximately 70-80 percent of them committing a new crime during the average follow-up period of well over 2 years. Most newly committed crimes were non-violent in nature, but a relatively high percentage of the juveniles who did fall back into their delinquent behavior also committed one or more violent crimes. This is not surprising, considering that data on the youth reentry population shows that almost half of the crimes that are newly committed are at the same severity level as the previous offense, and some subsequent offenses were even more serious in nature (Snyder & Sickmund, 2006). The general results are in line with other studies, showing that recidivism rates are essentially unaffected by aftercare programs that aim to achieve successful reintegration by providing intensive reentry services and follow-up care (Abrams, Shannon, & Sangalang, 2008; Frederick & Roy, 2003; Rowland, 2007; Wiebush et al., 2005).

Moderator analyses were conducted to determine whether NPAP was more or less successful with certain types of offenders. The results showed that none of the included variables were associated with a greater likelihood of success or fail-
ure in the NPAP group. One of the risk factors predictive of overall recidivism, violent and non-violent recidivism was the number of prior offenses youths committed, but no differences were found between the intervention and control group youths. Youths with a longer criminal record seem to be more vulnerable to recidivism, which is in line with Moffitt’s well-known dual taxonomy of offending behavior, indicating that Life-Course-Persistent offenders begin to behave antisocially early on in life and continue with their criminal behavior into adulthood (Moffitt, 1993). Moreover, the moderator analyses showed that age at first arrest is a risk factor for non-violent re-offending, which is consistent with prior research showing that the number of prior offenses and age of onset are predictive of further offending (Benda, Corwyn, & Toombs, 2001; Gendreau, Little, & Goggin, 1996; Loeber, & Farrington, 1998; Moffitt, Caspi, Harrington, & Milne, 2002). Given that these are static risk factors, it is impossible to offset their detrimental effects by means of aftercare services for juvenile and young adult offenders. However, determining and strengthening the possible impact of protective factors might compensate for these detrimental effects (Ward, Yates, & Willis, 2012).

Similar to results from a meta-analysis by Wilson and colleagues (2003), no differences in treatment effects were found for the various ethnic groups, indicating that NPAP does not seem to be culturally biased, because there was no differential impact on Dutch and ethnic minority youths.

Although no evidence was found to support the claim that NPAP can significantly reduce recidivism, the results of the study should not be interpreted as showing that NPAP is ineffective, that is, one should be careful not to accept the null hypothesis (Weisburd, Lum, & Yang, 2003). There are several explanations why NPAP was not more effective than TAU. Firstly, the majority of the control group youths received treatment as usual, which consisted of some form of (often theoretically based) intervention that in many ways was comparable to NPAP with respect to targeting similar criminogenic needs. Moreover, a recent meta-analytic study showed that aftercare yielded more positive effects if the control group in the study did not receive any intervention compared to a control group receiving ‘care as usual’ (James et al., 2013). Therefore, one can expect to find smaller effect sizes in the current study where the majority of the control group did receive some form of aftercare, and this may partially account for not finding any differences between the intervention and control group (Lösel, 2012).

Furthermore, it is widely recognized that program integrity is an important factor that can influence the effectiveness of an intervention (Andrews & Dowden, 2005; Gendreau, Coggin, & Smith, 1999). A recent meta-analytic review on aftercare programs has shown that aftercare can be effective if it is well-implemented, intensive in nature and aimed at high-risk youths (James et al., 2013).
Therefore, another possible explanation of not finding any evidence to support the claim that NPAP is an effective aftercare program is that NPAP was not carried out as intended. Unfortunately, insufficient information on participants’ trajectories was available for all NPAP participants who were included in the study to establish the level of treatment adherence (i.e., if NPAP was carried out according to the manual and provided to the stated target group). The involved youth care organizations documented too little information on all of the participants to establish, for example, the intensity of the provided aftercare, and if the inclusion criteria of recidivism risk and IQ level were met to allow for analyses of these possible moderators of the effectiveness. Hence, program integrity should be included in future research.

Thirdly, it is also possible that certain parts or specific aspects of the intervention are more effective than others. The inactive or redundant treatment elements could influence the effectiveness of NPAP and be responsible for not finding any significant differences between the NPAP and treatment as usual groups. Providing that sufficient information on the NPAP trajectories and provided treatment and guidance becomes available, future mediator analyses could identify possible mechanisms of change. Following this, NPAP could be redefined according to the active therapeutic components to enhance the overall effectiveness and subsequent recidivism reduction (Kraemer et al., 2002).

Finally, finding statistically significant differences could partly be due to the nature of the study design. Research has shown that the most rigorous designs, such as used in the present study, usually show smaller effects or no effects (Welsh, Peel, Farrington, Elffers, & Braga, 2011). This is attributable to limited internal validity in weaker research designs, where the comparison group tends to differ from the intervention group, more than is possible in an RCT, such as ours, where differences are controlled for by means of randomization. The latter results in a smaller likelihood to find differences in recidivism rates (Lipsey & Wilson, 1998; Weisburd, Lum & Petrosino, 2001).

The aim of NPAP is to promote desistance from crime in juvenile and young adult offenders. They are at the highest risk for reoffending and difficult to treat due to fluctuating motivation, a long history of criminal offending and ingrained antisocial behavior patterns. For this target group, NPAP often seems a matter of last resort with only limited chance of success (Skeem, Polashek, & Manchak, 2009). It seems unrealistic for those targeted to stop offending and never to commit another offense in their lifetime, considering the level of offending of our sample prior to referral to aftercare, both in frequency and in severity. Furthermore, many of the NPAP youths lacked the soft ‘people’ skills needed to land and maintain a steady job, such as good communication skills and positive work ethics, notwithstanding relentless efforts of their youth care workers to im-
prove their prosocial skills (James et al., 2014). Finally, they all carried the stigma of a (long) criminal record, making it even more difficult to succeed in building a stable and crime-free life for themselves.

This study has a number of strengths compared to many other studies on the effectiveness of aftercare programs. One of the strengths of the study is the strong research design, since the applied randomized controlled trial is the gold standard to study the effectiveness of interventions and the most robust in ruling out other factors of influence on the observed outcomes. Only few experimental studies have been carried out to establish the effectiveness of aftercare (James et al., 2013), and this study is the first RCT on an aftercare program for juvenile and young adult offenders in The Netherlands. Another strength of the study is its ecological validity. Since the study was carried out as a field experiment in the community, the results can be generalized and directly applied to practice and policy, aimed at improving aftercare services provided for juvenile and young adult offenders (Cape & Barkham, 2002; Van der Put et al., 2013).

However, there are also several limitations to be noted: firstly, the present study solely relies on official records to measure recidivism. Although official records offer a useful method of estimating recidivism, they always leave a ‘dark number’, and reflect only a part of the real level of reoffending. Criminal behavior that is not traced and brought before court is beyond the reach of the judicial data. In theory, this could mean that other results could have been found if a larger part of the offenses were included in the analyses. Self-report measures could be another source of information and a valuable contribution. On the other hand, self-report measures have a limitation of their own, since offenders are less likely to report the more serious offences they have committed and problems with memory and recall may arise (Piquero, Schubert, & Brame, 2014). Moreover, our sample comprised serious offenders who may be perceived as more inclined to misrepresent the truth (Spidel, Hugues, Greaves, & Yuille, 2011), making the official records used on the current study a trustworthier source of information. Furthermore, program integrity was not included in the study as a moderator, hence it cannot be ruled out that the results were influenced by the way NPAP was carried out.

Based on this study, some recommendations for further research can be made. Firstly, for policy makers and politicians to focus on recidivism alone as a measure of successful reintegration is not fair. It gives a limited view, because recidivism does not provide any indication of how well an individual is adjusting to community life. Therefore, it may be useful in future studies to also incorporate other measures of successful reintegration, such as housing stability, finance and social support, and finding and maintaining employment or an education (Graffam, Shinkfield, & Lavelle, 2014; Shinkfield, & Graffam, 2009). Further-
more, as others suggest (Asscher et al., 2007; Lösel, 2012), cost effectiveness and cost benefit analyses should become part of the outcome evaluation of NPAP. Cost-benefit analyses are an important, but often-neglected aspect of program evaluations, whereas successful aftercare interventions cannot only reduce harm done to society, but also save the society money in the long run.

In sum, the largely null effects found in the present experimental study indicate that perhaps stronger aftercare programs and components should be developed, possibly more specifically focused on other risk factors than cognitions and with a higher level of program integrity, to establish a significant decrease in criminal offending behavior for these vulnerable youths in transition from adolescence to adulthood and from a life of crime to one as law-abiding citizens. On the other hand, not finding superior effects of NPAP compared to treatment as usual does not indicate that NPAP does not work at all. NPAP is a theory driven and multifaceted aftercare intervention developed to deal with the most problematic juvenile and young adult offenders in the Dutch criminal justice system who, typically, experience wide-ranging challenges to reintegration. Much remains to be learned on how to successfully reduce recidivism among this group of serious young offenders. Our RCT contributes to the small, but important body of empirical work, and certainly bears further investigation into the effectiveness of aftercare programs.
Table 1. Descriptives of the intervention (NPAP) group and control (TAU) group at pretest

<table>
<thead>
<tr>
<th>Category</th>
<th>NPAP</th>
<th>TAU</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>77</td>
<td>97.5</td>
</tr>
<tr>
<td>Age (years)</td>
<td>Mean (SD)</td>
<td>19.90 (2.58)</td>
<td>19.50 (2.36)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Dutch</td>
<td>8</td>
<td>10.1</td>
</tr>
<tr>
<td>Non-western immigrant</td>
<td>37</td>
<td>46.8</td>
<td>33</td>
</tr>
<tr>
<td>Caribbean</td>
<td>16</td>
<td>20.3</td>
<td>25</td>
</tr>
<tr>
<td>Other</td>
<td>18</td>
<td>22.8</td>
<td>8</td>
</tr>
<tr>
<td>NPAP provided on a voluntary or mandatory basis</td>
<td>Voluntary</td>
<td>48</td>
<td>60.8</td>
</tr>
<tr>
<td>Age of first police contact</td>
<td>Mean (SD)</td>
<td>15.36 (1.92)</td>
<td>15.44 (1.89)</td>
</tr>
<tr>
<td>Number of prior offenses</td>
<td>Mean (SD)</td>
<td>5.77 (3.49)</td>
<td>5.17 (3.98)</td>
</tr>
<tr>
<td>Number of prior violent offenses</td>
<td>Mean (SD)</td>
<td>1.85 (1.42)</td>
<td>1.87 (1.80)</td>
</tr>
<tr>
<td>Number of prior non-violent offenses</td>
<td>Mean (SD)</td>
<td>3.90 (3.20)</td>
<td>3.28 (3.32)</td>
</tr>
</tbody>
</table>

Note: NPAP = New Perspectives Aftercare Program; TAU = treatment as usual. * p < .05, ** p < .01.

Table 2. Official judicial reconviction data at follow-up for intervention (NPAP) group and control (TAU) group

<table>
<thead>
<tr>
<th>Category</th>
<th>NPAP</th>
<th>TAU</th>
<th>F</th>
<th>d [ 95% CI]¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of reconvictions</td>
<td>2.51 (2.8)</td>
<td>2.40 (2.7)</td>
<td>.121</td>
<td>.06 [-.29, .42]</td>
</tr>
<tr>
<td>Number of violent reconvictions</td>
<td>.90 (1.12)</td>
<td>.91 (1.12)</td>
<td>.002</td>
<td>.01 [-.35, .37]</td>
</tr>
<tr>
<td>Number of non-violent reconvictions</td>
<td>1.62 (1.84)</td>
<td>1.51 (1.34)</td>
<td>.138</td>
<td>.07 [-.29, .43]</td>
</tr>
<tr>
<td>Time to first reconviction (in days)</td>
<td>401.04 (352.83)</td>
<td>458.12 (390.07)</td>
<td>.946</td>
<td>.18 [-.18, .54]</td>
</tr>
<tr>
<td>Time to first violent reconviction (in days)</td>
<td>591.39 (362.92)</td>
<td>612.70 (377.43)</td>
<td>.133</td>
<td>.07 [-.29, .43]</td>
</tr>
<tr>
<td>Time to first non-violent reconviction (in days)</td>
<td>576.01 (423.02)</td>
<td>577.57 (392.16)</td>
<td>.001</td>
<td>.01 [-.35, .36]</td>
</tr>
</tbody>
</table>

Note: CI= confidence interval; NPAP = New Perspectives Aftercare Program; TAU = treatment as usual. * N only for those juveniles who did recidivate. * p < .05, ** p < .01.

18. Second generation non-western migrant workers
19. Second generation from former Dutch Caribbean colonies
20. Offense related data based on official judicial data from the Judicial Information Service (JustID)
Figure 1. Flow diagram of participants in the study of the effectiveness of NPAP based on recidivism outcomes.

Assessed for eligibility and referred to primary researcher for inclusion ($n = 183$)

Enrollment

- Excluded ($n = 20$)
  - Declined to participate ($n = 1$)
  - Did not meet inclusion criteria ($n = 1$)

Randomized ($n = 163$)

Allocated to NPAP ($n = 79$)
- Received allocated intervention ($n = 62$)
- Did not receive allocated intervention ($n = 17$)

Allocated to TAU ($n = 84$)
- Received allocated intervention ($n = 56$)
- Did not receive allocated intervention ($n = 25$)

 analyzes ($n = 79$)
- Excluded from analysis ($n = 0$)

Analyses ($n = 82$)
- Excluded from analysis ($n = 2$)

Note. NPAP = New Perspectives Aftercare Program and TAU = treatment as usual. This flowchart is an adaptation of the flowchart offered by the CONSORT Group (Altman et al., 2001; Moher, Schulz, & Altman, 2001).
Figure 2: Survival curve for overall recidivism for NPAP and TAU groups separately
Figure 3. Survival curve for violent recidivism for NPAP and TAU groups separately.
Figure 4. Survival curve for non-violent recidivism for NPAP and TAU groups separately