The outcome of non-residential youth care compared to residential youth care: A multilevel meta-analysis*

Raymond V. Gutterswijk,a,b⁎, Chris H.Z. Kuiperb,c, Navisha Lautanb, Elsemieke G. Kunstd, Frank C.P. van der Horsta, Geert Jan J.M. Stams,e, Peter Prinziea

a Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, the Netherlands
b Horizon Youth Care and Special Education, Mozartlaan 150, 3055 KM Rotterdam, the Netherlands
c Social Work and Applied Psychology, University of Applied Sciences, Zernikedreef 11, 2333 CK Leiden, the Netherlands
d Criminal Law and Criminology, VU University, De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands
e Forensic Child and Youth Care Sciences, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Foster care treatment
Home-based treatment
Residential youth care
Meta-analysis
Behavioral problems
Effectiveness

ABSTRACT

Objective: This multilevel meta-analysis compared the outcomes of Treatment Foster Care Oregon for Adolescents (TFCO-A) and home-based treatment programs (HBT) with residential youth care for children and youth aged 0 to 23 years.

Methods: A total of 145 effect sizes for different types of behavioral problems were derived from 24 controlled studies (n = 16,943 participants). A three-level random-effects meta-analysis was conducted.

Results: We found a small statistically significant overall effect (d = 0.21), 95% CI [0.090-0.338], which indicated that non-residential youth care was slightly more effective than residential youth care. However, moderator analysis revealed that TFCO-A yielded a larger effect size (d = 0.36) than HBT (d = 0.08).

Conclusions: Our findings suggest that youth treated in treatment foster care have better outcomes than youth in residential care, which is not true for children who are treated at home. Therefore, in case of out-of-home placement treatment foster care should be the first option. Given that residential care has no additional value for youth who are treated at home, and often sets limits to juveniles’ needs for self-determination, residential care seems an option if TFCO-A is not available and living at home is no longer possible because the child’s (immediate) safety is at stake.

1. Introduction

There is an ongoing debate on how to effectively treat youth with complex needs who are at risk for out-of-home placement, especially (therapeutic) round-the-clock care in residential settings (Whittaker et al., 2016). These youths experience severe problems in behavioral functioning at home, in school and during leisure activities (Attar-Schwartz, 2009; Eliak et al., 2017; Frensch, & Cameron, 2002; Leloux-Opmeer, Kuiper, Swaab, & Scholte, 2017; Martín, González-García, Del Valle, & Bravo, 2018). The most common reason for referral of a youth to residential care is the presence of serious parenting and behavioral problems (Ainsworth, 2017; Bruning & De Jong-De Kruifj, 2015; Leloux-Opmeer et al., 2017; Martín et al., 2018). Alternatively, youth may be placed in a (forensic) secure residential institution after having committed a crime. Residential youth care, however, is the most intensive and most expensive type of youth care, which substantially restricts autonomy of children and adolescents and deprives them of family life, which is particularly undesirable for youth who are placed in residential care because home-based treatment or foster care is not available (Busschers & Konijn, 2019) or due to long-term under-treatment of severe behavior problems (Broeders, Van der Helm, & Stams, 2015). Residential youth care may even cause harm when youths are exposed to institutional repression or negative peer influences (De Valk, Kuiper, Van der Helm, Maas, & Stams, 2016; Dishion, McCord, & Poulin, 1999). Residential youth care is therefore mostly seen as a ‘placement of last resort’.

The last decade, serious doubts have been raised about the effectiveness and appropriateness of residential youth care (Souverein, Van

---

* Funding: this research was funded by Reformed Civil Orphanage Rotterdam.

⁎ Corresponding author at: Horizon Youth Care and Special Education, Mozartlaan 150, 3055 KM Rotterdam, the Netherlands.

E-mail addresses: gutterswijk@essb.eur.nl (R.V. Gutterswijk), kuiper.c@hsleiden.nl (C.H.Z. Kuiper), vanderhorst@essb.eur.nl (F.C.P. van der Horst), g.j.j.m.stams@uva.nl (G.J.J.M. Stams), prinzie@essb.eur.nl (P. Prinzie).

https://doi.org/10.1016/j.childyouth.2020.104950

Received 18 November 2019; Received in revised form 19 March 2020; Accepted 19 March 2020

Available online 20 March 2020
The last decades several programs have been developed as alternative to residential or institutional youth care. Treatment Foster Care Oregon for Adolescents (TFCO-A), formerly known as multi treatment foster care (MTFC), aims to reduce deviant behavior (Bergström & Höjman, 2016; Chamberlain, Leve, & DeGarmo, 2007; Sinclair et al., 2016), and delinquent activity in youth (Chamberlain & Reid, 1998). TFCO-A also aims to reinforce prosocial behavior (Bergström & Höjman, 2016; Chamberlain et al., 2007) by encouraging participation in structured social activities, social skills training, and fostering good relationships with parents and peers (Fisher & Chamberlain, 2000). TFCO-A consists of an out-of-home placement in a professionally trained foster family for 6 to 9 months. In addition, a clinical team is formed around the youth and his or her birth family (Fisher & Chamberlain, 2000; Westermark, Hansson, & Olsson, 2011). The youth is offered a therapeutic and structured living environment, where supervision, boundary setting, and supporting relationships are important. The most important difference with residential care is that the youth lives within a family context, and mostly receives education at a regular school (see e.g., The California Evidence-based Clearinghouse, 2018).

Another form of non-residential care, Home-based treatment (HBT), is offered to youth living at home. HBT targets youths with serious emotional and behavioral problems who are at risk of being placed out-of-home or return home from an out-of-home placement (Mattejat, Hirt, Wilken, Schmidt, & Remschmidt, 2001). By implementing HBT, organizations aim to improve the overall well-being of the family and reduce problems affecting the family (Preyde et al., 2011b). Recently, four types of HBT which are highly comparable in used mechanisms and techniques and in treatment effects have been compared with residential care (Van der Pol et al., 2017). These types of HBT are Intensive Home-Based Treatment (IHBT), Multidimensional Family Therapy (MDFT), Functional Family Therapy (FFT) and Multi-Systemic Therapy (MST).

Intensive Home-Based Treatment (IHBT) is defined as all out-patient youth care for more than one hour per week. IHBT promotes positive development and adequate family functioning. IHBT addresses mental health issues and is available ‘around the clock’. IHBT is offered both individually and systemically (Moffett, Brotnow, Patel, Adnopoz, & Woolstone, 2017). In their meta-analysis on the outcomes of wrap-around care, Suter and Bruns (2009) found a small statistically significant effect (Cohen’s d = 0.33) on youths’ mental health and overall functioning.

Multidimensional Family Therapy (MDFT) offers help to youth with multiple problems behavior. The purpose of MDFT is to make youths’ problematic behavior disappear or decrease and improve youths’ functioning within the family, in school or work, and in daily life. The therapists using MDFT involve family, friends, school, work and promote leisure activities. Furthermore, the meta-analysis performed by Van der Pol and colleagues (2017) found a small, but significant, overall effect size (d = 0.24) of MDFT compared to other therapies on various behavioral outcomes.

Functional Family Therapy (FFT) is aimed at enabling the family to resolve problems themselves and to deal with setbacks, and works with 11- to 18-year old youths who have been referred for behavioral or emotional problems. FFT can be offered at home, in school or in a mental health facility (Robbins, Alexander, Turner, & Hollimon, 2016). Hartnett, Carr, Hamilton, and O’Reilly (2017) performed a meta-analysis on the effects of FFT on drug use, recidivism, family adjustment and behavioral problems, and found small statistically significant treatment effects compared to untreated control groups (Cohen’s d = 0.48) and alternative treatments (Cohen’s d = 0.35).

Multi-Systemic Therapy (MST) strongly focuses on the network at large, including the school, peers, and sports clubs (Henggeler, 2011), improving communication, parenting skills, peer relations, school performance, and social networks (Little, Popa, & Forsythe, 2005). Van der Stouwe, Asscher, Stams, Deković, and Van der Laan (2014) conducted a meta-analysis and reported that MST produced small statistically significant positive effects on delinquency (Cohen’s d = 0.20), psychopathology (Cohen’s d = 0.27), substance abuse (Cohen’s d = 0.29), family factors (Cohen’s d = 0.14), out-of-home placements (Cohen’s d = 0.27), and peer factors (Cohen’s d = 0.21).

1.3. The present study

In this multilevel meta-analysis, we compare the outcomes of two well-researched types of non-residential interventions that target youth with complex problems at risk for out-of-home placement – Treatment Foster Care Oregon for Adolescents (TFCO-A) and home-based treatment (HBT) – with the outcomes of residential youth care. We only include controlled studies comparing TFCO-A or HBT to residential youth care, reporting on internalizing, externalizing, and total
behavioral problems, substance misuse and delinquency for children and adolescents aged 0 to 23 years. This meta-analysis is innovative because, in contrast with the meta-analyses of Strijbosch and colleagues (2015) and De Swart and colleagues (2012), in the present meta-analysis (1) initial differences between youth receiving residential and non-residential care are accounted for by means of study design (matching or randomization) and control for pre-test differences; and (2) not only differences in effect sizes between studies, but also within studies are taken into account by means of recent developed meta-analytic techniques (Assink & Wibbelink, 2016). We thus gain knowledge on differences between the outcomes of residential and non-residential youth care in children and adolescents with comparable problems and supposedly risks for out-of-home placement, and the conditions under which these outcomes differ. This knowledge can be used to inform clinical practice and policies on the delivery of residential and non-residential youth care.

Overall and based on the literature, we hypothesize that non-residential youth care will produce more favorable outcomes than residential care, because residential care may have a negative effect on the developmental possibilities and treatment motivation of youth by: (1) not meeting the fundamental requirements for self-determination (i.e., competence development, contact and autonomy); (2) the association with deviant peers and deviancy training; (3) high risks for institutional repression; (4) the unavailability of evidence-based manualized treatment; (5) the cut off from primary supportive attachment-based relationships; and (6) problems in establishing supportive (therapeutic) youth-staff relationships (De Valk et al. 2016; Souverein, Van der Helm, & Stams, 2013; Van der Helm, Kuiper, & Stams, 2018).

In moderator analyses we examine the degree to which the overall effect size for differences in youth outcomes between residential and non-residential care is affected by sample characteristics (e.g., mean age, sex and ethnicity), methodological characteristics (e.g., study design, quality of the study and control for pre-test), and study characteristics (e.g., impact factor, type of intervention and year of publication). These moderators are included in meta-analyses as a rule, because they control for methodological influences and publication characteristics or concern generalizability of results.

In previous meta-analyses comparing effectiveness of non-residential and residential care, moderation effects were found for gender (Strijbosch et al., 2015), but not for age (De Swart et al., 2012; Strijbosch et al., 2015), showing that studies with a high percentage of females were associated with smaller effect sizes. In the meta-analysis by Strijbosch and colleagues, study design (i.e., randomized controlled trial, matched or non-matched control group) was a significant moderator, showing that matched studies yielded better outcomes for youth in non-residential care, whereas non-matched studies showed better outcomes for youth in residential care. However, study design was non-significant as a moderator in the meta-analysis by De Swart et al.. In
contrast to the meta-analysis of De Swart and colleagues (2012), the meta-analysis of Strijbosch and colleagues (2015) showed a significant moderator effect for type of outcome. Year of publication was a significant moderator in both meta-analyses indicating that earlier published studies yielded greater effect sizes. Finally, De Swart et al. found a moderator effect for the type of intervention, with only positive effects for cognitive behavior therapy.

2. Methods

2.1. Study selection

We searched for studies on residential youth care in several electronic databases: Pubmed, SAGE Journals, ScienceDirect, SpringerLink, Wiley Online Library, MEDLINE, Web of Science, CINAHL, PsycINFO, Cochrane Library, Campbell Library, Proquest and Google Scholar. To cover the terms child, residential care, antisocial behavior and treatment effect, we used the following set of keywords: youth, child*, adolescent*, boy*, girl*, juvenile*, residential care, residential homes, institutional care, group care, group homes, problem behavior*, behavio* problems, aggres*, violen*, criminal behavio*, antisocial behavio*, externalizing, delinquen*, internalizing, anxiety, depression, effect* and comparison. In addition, we inspected the reference lists of the studies we included in this meta-analysis. Finally, two researchers searched independently the indexes of the most relevant journals. The final search was performed on September 4, 2019.

2.2. Inclusion criteria

Studies that met the following conditions were included: (1) the (quasi-) experimental group or the control group received home-based treatment or Treatment Foster Care Oregon for Adolescents, (2) the other group received residential care, (3) the studies provided at least post-test scores or follow-up scores for both groups and (4) were written in English or Dutch. We included a total of 24 studies (N = 16,943) (see Fig. 1). The literature search was performed by three researchers. When in doubt whether a study did meet the inclusion criteria, the three researchers discussed what to do until consensus was reached. No unpublished relevant studies were found. The reason to exclude studies on the basis of the full-text was mainly because the control group or outcome measures did not meet the inclusion criteria.

Two researchers coded all available outcome variables of the studies we included in the meta-analysis. The first five studies were coded independently by both researchers. Inter-rater reliability was analyzed by calculating Kappa for categorical variables and intraclass correlation (ICC) for variables at the interval and ratio level. This inter-rater reliability was moderate to nearly perfect, according to the guidelines by Landis and Koch (1977). Our Kappa’s ranged from 0.70 to 1.00 and intraclass correlations from 0.99 to 1.00. In one case, the Kappa proved to be insufficient, yielding a score of 0.54 which problem was resolved through further discussion until consensus was reached. A limitation of ICC is that it does not include missing values in the analysis. This is an important improvement, because commonly used meta-analytic methods assume independency of effect sizes, whereas this usually is not the case. The method also allows for the use of multiple effect sizes (within studies) from the same sample (Assink & Wibbelink, 2016). In a funnel plot, each effect size is plotted on the horizontal axis against its sample size, standard error or precision on the vertical axis. This distribution is shaped as a funnel if no publication bias is present. A violation of funnel plot asymmetry indicates publication bias. By regressing the standard normal deviate, defined as the effect size divided by its standard error, against the estimate’s precision, funnel plot asymmetry can be tested. If there is asymmetry, the regression line does not run through the origin and the intercept significantly deviates from zero (Duval & Tweedie, 2000).

2.3. Publication bias funnel plot

Studies reporting strong significant results are more likely to be published in peer-reviewed journals. Studies reporting less strong or no statistically significant results are therefore harder to find. To examine file drawer bias, a funnel plot of the distribution of effect sizes can be used (Rosenthal & Hirstein, 1979). In a funnel plot, each effect size is plotted on the horizontal axis against its sample size, standard error or precision on the vertical axis. This distribution is shaped as a funnel if no publication bias is present. A violation of funnel plot symmetry indicates publication bias. By regressing the standard normal deviate, defined as the effect size divided by its standard error, against the estimate’s precision, funnel plot asymmetry can be tested. If there is asymmetry, the regression line does not run through the origin and the intercept significantly deviates from zero (Duval & Tweedie, 2000).

2.4. Analyses of effect sizes

To analyze our data, we used a random effects model (Sánchez-Meca & Marín-Martínez, 2008). We calculated Cohen’s d for group comparisons, using the Practical Meta-Analysis Effect Size Calculator developed by Wilson (2001). Cohen’s d was calculated by using means and standard deviations, proportions and t-, F-, χ²-, p-values. For 128 of 145 effect sizes it was possible to control for pre-test scores.

The multilevel meta-analysis was conducted in R (version 3.5.1), using the metaphor-package (Viechtbauer, 2017). In a three-level meta-analysis, variance at three different levels is analyzed: (1) sample variance, (2) variance between effect sizes within studies, and (3) variance among effect sizes between studies (Assink et al., 2018; Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2013). The multilevel technique allows not only to calculate an overall effect size, but if significant variance is present at level 2 and/or 3, moderation by sample, methodological and/or study characteristics can be examined. This is an important improvement, because commonly used meta-analytic methods assume independency of effect sizes, whereas this usually is not the case. The method also allows for the use of multiple effect sizes (within studies) from the same sample (Assink & Wibbelink, 2016). Moderator analyses were also performed using R (Viechtbauer, 2017).

3. Results

3.1. Study characteristics

This meta-analysis included k = 24 primary studies from which u = 145 effect sizes were extracted (see Appendix A). On average, 6.04 effect sizes were extracted from each included study (SD = 5.06; range = 1–20). The studies were published between 1992 and 2018, and the median year was 2005. Almost all studies were conducted in North-America (k = 22), with only two European studies.
3.2. Overall effect size

The estimated overall effect of non-residential care on behavioral problems, compared to residential youth care, was $d = 0.21$, $p \leq 0.001$, 95% CI [0.090, 0.338] (see Table 1), indicating that youth in non-residential care showed statistically significantly better outcomes than youth in residential care. According to Rice and Harris (2005) this is a small effect. Significant level 2 and level 3 variance was found. This significant variance implies substantial variability in effect sizes extracted from the same study (level 2) and from different studies (level 3) (see Table 1). As presented in Table 1 about 62% of total variance could be explained by within-study differences in effect sizes (level 2) and about 32% by between-study differences in effect sizes (level 3). Therefore, moderator analyses were conducted to explore whether moderators could explain the level 2 and 3 heterogeneity.

Furthermore, a visual inspection of the funnel plot did not lead to a suspicion of publication bias, which was confirmed by the trim-and-fill analysis in R, which revealed that no effect sizes had to be imputed at the left or right side of the funnel (see Fig. 2).

3.3. Moderator analyses

In Table 2, the results of within-study moderator analyses are presented. In Table 3, the between-study moderator analyses are presented. The moderators are classified into ‘sample characteristics’, ‘methodological characteristics’ and, ‘study characteristics’.

### 3.3.1. Within-study: Methodological characteristics

We found no moderating effect of the measured outcomes (total behavior problems, internalizing behavior problems, externalizing behavior problems, substance abuse or delinquency), type of measure (questionnaire, interview or official registration), information source (child, parent or other), measurement moment (post-test or follow-up) or number of months before follow-up.

### 3.3.2. Between-study: Sample characteristics

We found no moderating effect of percentage male, age, percentage Caucasian white, percentage African black or percentage Hispanic.

### 3.3.3. Between-study: Methodological characteristics

Moderator analysis revealed a statistically significant moderating effect for type of intervention (see Table 3). TFCO-A yielded a larger effect ($d = 0.36$) on behavioral problems than HBT ($d = 0.08$), indicating TFCO-A to be slightly more effective than residential youth care, whereas HBT proved to be equally effective. No moderating effects were found for year of publication, intention to treat, study design, study quality, author, control for pretest or impact factor of the journal the study was published in.

4. Discussion

The main aim of this meta-analysis was to examine the outcomes of non-residential care compared to residential care for youth aged 0 to 23 years, which revealed that TFCO-A yielded a larger effect size (Cohen’s $d = 0.36$) than HBT (Cohen’s $d = 0.08$). These findings

---

**Table 1**

| Outcome          | $k$ | #ES | Mean d | 95% CI | Sig. mean d ($p$) | $\sigma^2_{level\ 2}$ | $\sigma^2_{level\ 3}$ | % Var. Level 1 | % Var. Level 2 | % Var. Level 3 |
|------------------|-----|-----|--------|--------|-------------------|----------------------|----------------------|----------------|----------------|----------------|
| Child-outcomes   | 24  | 145 | 0.21   | 0.00; 0.34 | 0.00 ***          | 0.12 ***             | 0.06 ***             | 6.61           | 61.60          | 31.79          |

Note. Child outcomes = internalizing problem behavior, externalizing problem behavior, substance use, delinquency and total problems; $k =$ number of studies; #ES = number of effect sizes; mean $d =$ mean effect size (Cohen’s $d$); CI = confidence interval; $\sigma^2_{level\ 2}$ = variance between effect sizes extracted from the same study; $\sigma^2_{level\ 3}$ = variance between studies; % Var = percentage of variance distributed.

*** $p \leq 0.001$.

Fig. 2. Trim and fill plot for all effect sizes. Note. A contour enhanced funnel plot with Cohen’s $d$ on the X axis and standard error on the Y axis. The black dots represent the extracted effect sizes. If there were any imputed effect sizes, they would be represented by white dots. The solid vertical line represents the overall effect size.
indicate that treatment foster care was slightly more effective than residential care, whereas home-based care proved to be equally effective. The positive effects of non-residential (foster) care of this multilevel meta-analysis are largely in line with findings of the meta-analyses by De Swart and colleagues (2012) and Strijbosch and colleagues (2015), who found small (Cohen’s $d = 0.20$) and small-to-medium (Cohen’s $d = 0.34$) effects, respectively, favoring non-residential over residential youth care. Therefore, the combined findings of previous meta-analyses.

### Table 2
Within-study moderators of the effectiveness of non-residential care: assessment of outcomes.

| Moderator variable                                  | $k$ | #ES | $b_0/d$ | $t_0$ | $b_1$ | $t_1$ | $F(df_1, df_2)$ |
|-----------------------------------------------------|-----|-----|---------|-------|-------|-------|-----------------|
| **Methodological characteristics**                  |     |     |         |       |       |       |                 |
| **Measured outcomes**                                |     |     |         |       |       |       |                 |
| Total behavior problems                              | 10  | 24  | 0.04    | 0.35  |       |       | $F(4, 140) = 1.67$ |
| Internalizing behavior problems                      | 12  | 45  | 0.21    | 2.53* | 0.17  | 1.62  |                 |
| Substance abuse                                      | 5   | 13  | 0.12    | 0.88  | 0.09  | 0.53  |                 |
| Delinquency                                          | 14  | 34  | 0.35    | 3.78***| 0.31  | 2.37  |                 |
| Externalizing behavior problems                      | 11  | 29  | 0.14    | 1.53  | 0.10  | 0.95  |                 |
| **Approaches to Outcome Measurement**                |     |     |         |       |       |       |                 |
| Type of Measure                                      |     |     |         |       |       |       |                 |
| Questionnaires                                       | 17  | 93  | 0.16    | 2.18* |       |       | $F(2, 136) = 2.24$ |
| Official registration                                | 11  | 21  | 0.42    | 3.69***| 0.26  | 2.08  |                 |
| **Time of measurement**                              |     |     |         |       |       |       |                 |
| Post-test                                            | 9   | 48  | 0.20    | 2.80**| 0.05  | 0.51  |                 |
| Follow-up months                                     | 18  | 103 | 0.17    | 3.33***| 0.00  | 0.76  | $F(1, 101) = 0.57$ |

*Note. $k$ = number of independent studies; #ES = number of effect sizes; $b_0/mean r = intercept/mean effect size (r); t_0 = difference in mean $r$ with zero; $b_1 = estimated regression coefficient; t_1 = difference in mean $r$ with reference category; $F(df_1, df_2) = omnibus test; (RC) = reference category. + $p < .10; * $p < .05; ** $p < .01; *** $p < .001.

To illustrate, Table 2 displays within-study moderators of the effectiveness of non-residential care: assessment of outcomes. The table includes various categories such as methodological characteristics, type of measure, informant, time of measurement, and study characteristics. Each category is further divided into specific subcategories with corresponding values for $k$, #ES, $b_0/d$, $t_0$, $b_1$, $t_1$, and $F(df_1, df_2)$.

### Table 3
Between-study moderators of the effectiveness of non-residential care: child and methodological characteristics.

| Moderator variable                                  | $k$ | #ES | $b_0/d$ | $t_0$ | $b_1$ | $t_1$ | $F(df_1, df_2)$ |
|-----------------------------------------------------|-----|-----|---------|-------|-------|-------|-----------------|
| **Sample characteristics**                           |     |     |         |       |       |       |                 |
| Sex                                                  | 24  | 145 | 0.20    | 3.18**| −0.00 | −0.99 | $F(1, 143) = 0.99$ |
| Age                                                  | 21  | 140 | 0.17    | 3.51***| 0.02  | 0.76  | $F(1, 137) = 0.58$ |
| Ethnicity                                            |     |     |         |       |       |       |                 |
| Percentage Caucasian White                           | 17  | 94  | 0.25    | 3.17**| 0.00  | 1.17  | $F(1, 927) = 1.37$ |
| Percentage African Black                             | 17  | 87  | 0.27    | 3.50***| −0.00 | −1.30 | $F(1, 85) = 1.68$ |
| Percentage Hispanic                                  | 15  | 84  | 0.18    | 2.83*  | −0.00 | −1.26 | $F(1, 82) = 1.58$ |
| **Methodological Characteristics**                   |     |     |         |       |       |       |                 |
| Study design                                         |     |     |         |       |       |       |                 |
| RCT                                                  | 13  | 73  | 0.27    | 3.05**|       |       | $F(2, 142) = 0.73$ |
| Quasi experimental matched                           | 7   | 28  | 0.22    | 1.81  | −0.05 | −0.29 |                 |
| Quasi exp. non-matched                               | 4   | 44  | 0.07    | 0.46  | −0.20 | −1.21 |                 |
| Study quality                                        |     |     |         |       |       |       |                 |
| Strong                                               | 8   | 56  | 0.29    | 2.65**|       |       | $F(2, 137) = 0.38$ |
| Moderate                                             | 8   | 57  | 0.20    | 1.77  | −0.09 | −0.58 |                 |
| Weak                                                 | 8   | 32  | 0.15    | 1.25  | −0.14 | −0.85 |                 |
| Research group                                       |     |     |         |       |       |       |                 |
| Other group                                          | 11  | 62  | 0.22    | 2.59* |       |       | $F(3, 141) = 1.93$ |
| Group Preyde                                         | 4   | 44  | 0.01    | 0.05  | −0.21 | −1.43 |                 |
| Group Chamberlain                                    | 6   | 22  | 0.41    | 3.40***| 0.19  | 1.27  |                 |
| Group Henggeler                                      | 3   | 17  | 0.12    | 0.78  | −0.10 | −0.57 |                 |
| Control for pretest                                  |     |     |         |       |       |       |                 |
| Control for pretest                                  | 20  | 129 | 0.33    | 2.27* |       |       | $F(1, 143) = 0.74$ |
| No control for pretest                               | 5   | 16  | 0.19    | 2.79***| −0.14 | 0.86  |                 |
| Intention to treat                                   |     |     |         |       |       |       |                 |
| Completers                                          | 11  | 64  | 0.23    | 2.40* |       |       | $F(1, 143) = 0.02$ |
| Intention to treat                                   | 14  | 81  | 0.21    | 2.44* | −0.02 | −0.12 |                 |
| **Study characteristics**                            |     |     |         |       |       |       |                 |
| Type of experimental group                           |     |     |         |       |       |       |                 |
| Homebased Services                                   | 12  | 99  | 0.08    | 1.11  |       |       | $F(1, 143) = 6.10*$ |
| TFCA-A                                               | 12  | 46  | 0.36    | 4.28***| 0.28  | 2.47  |                 |
| Impact factor                                        | 21  | 113 | 0.25    | 3.14**| −0.01 | −0.19 | $F(1, 111) = 0.04$ |
| Year of publication                                  | 24  | 145 | 0.21    | 3.27***| −0.00 | −0.21 | $F(1, 143) = 0.04$ |

*Note. $k$ = number of independent studies; #ES = number of effect sizes; $b_0/mean r = intercept/mean effect size (r); t_0 = difference in mean $r$ with zero; $b_1 = estimated regression coefficient; t_1 = difference in mean $r$ with reference category; $F(df_1, df_2) = omnibus test; (RC) = reference category. + $p < .10; * $p < .05; ** $p < .01; *** $p < .001.
and the present meta-analysis suggest that treatment foster care should be preferred above residential youth care in case of out-of-home placement.

All moderators, except for type of intervention (TFCO-A or HBT), turned out to be statistically non-significant, which indicates that there was no difference in the effect of non-residential care compared to residential care for boys and girls, young children and adolescents, youth of different ethnic backgrounds and measured outcome. In line with our findings, De Swart and colleagues (2012) found no significant moderator effects for gender, age, ethnicity and type of measured outcome either. However, the findings of Strijbosch and colleagues (2015) differed from our findings in the sense that samples with a larger percentage of females yielded smaller effect sizes. Notably, Sawyer, Borduin, and Dopp (2015) found smaller effect sizes for samples with more boys in their meta-analysis of the long-term effects of prevention and treatment of youth with antisocial behavior. While several authors claim that boys and girls are in need of a different approach (Baker, Archer, & Curtis, 2005; Herman, 1997; Zahn, Day, Mihalic, & Tichavsky, 2009), our findings suggest that girls and boys benefit in a similar way from treatment foster care if compared to residential care, although boys and girls may be selected for similarity, such as similar problem behaviors.

The time of measurement yielded no moderating effect either, indicating that the difference in effect between non-residential and residential care is stable over time, similar to results of the meta-analysis by De Swart and colleagues (2012). In addition, we did not find a moderating effect of measured outcomes, whereas Strijbosch (2015) found that non-residential care was more effective than residential care in reducing delinquency, but not more effective in producing positive outcomes in other domains of youth functioning. Our study indicates that the more positive treatment effects of foster care, in particular TFCO-A, pertain to all outcomes, including both internalizing and externalizing problems.

4.1. Implications for clinical practice and future research

Results of this meta-analysis and those of De Swart and colleagues (2012) and Strijbosch and colleagues (2015) indicate that treating a youth through non-residential care has a more positive effect than treating the youth within residential care. The great advantage of non-residential youth care is that the youth lives within a family and the parents of the youth can more easily be involved in treatment (Fischer & Chamberlain, 2000; Mattejat et al., 2001), instead of reducing opportunities for contact with the family in residential care (James, 2017).

Another important advantage of non-residential care is that potential iatrogenic effects of residential care are avoided, although these negative effects have been contested in several studies (Huefner, Handwerk, Ringle, & Field, 2009; Huefner & Ringle, 2012; Lee & Thompson, 2009). There is some empirical evidence showing that working on a therapeutic residential group climate may neutralize possible iatrogenic effects (Stams & Van der Helm, 2017). These iatrogenic effects may in particular be caused by ‘deviancy training’. This means that deviant peers reinforce each other’s antisocial behaviors when care is provided to a group instead of individually (Dishion, Poulin, & Burriston, 2001; Weiss et al., 2005). Furthermore, residential youth care is a very intensive and expensive type of youth care, restricting youths’ autonomy (James, 2017; Knorth et al., 2007), and their need for self-determination (Van der Helm et al., 2018). In general, a trajectory in TFCO-A is less expensive than a placement in secure residential youth care. TFCO-A, however, is slightly more expensive than a placement in residential youth care if length of stay is comparable (Åström et al., 2019). This is why serious doubts have risen about the effectiveness and appropriateness of secure residential youth care (Souverein et al., 2013).

Although the outcomes for youth in non-residential care were only slightly better than those for youth in residential care, these small improvements can be meaningful in the long run. A small effect can be very important in some cases, especially if interventions target severe problems, in this particular case, severe behavioral problems in youth (Thompson, 2007). Notably, our research findings were based on multiple studies, accounting for both within and between study heterogeneity. The results of this meta-analysis indicate that for many youth non-residential care is the preferred option, especially TFCO-A, both in terms of achievement of therapeutic objectives and cost-effectiveness (McCartney & Rosenthal, 2000). However, we are aware that TFCO-A is only available to a limited extent. We therefore strongly recommend that the availability of TFCO-A be expanded, for example, by replacing a part of residential youth care by TFCO-A. Nevertheless, it cannot be ruled out that there still may be a specific group of the most troubled youths, such as adolescents with psychopathic traits (Asscher et al., 2011) or early onset conduct disorder (Wibbelink, Hoeve, Stams, & Oort, 2017), who are unsuitable for treatment at home or in foster care, and for whom residential youth care is the only viable option. Sometimes behavioral problems are so severe that youths are unmanageable within their own family or even in a foster family, with high risks of foster care placement instability (Konijn et al., 2019; Van den Bergh, Weterings, & Schoenmakers, 2011).

If we want to prevent youth from entering residential care, we need to find out when youth should still be referred to residential care or sentenced to detention, when no alternative sanction is allowed, and how alternative interventions can be developed for youth who depend on residential youth care because of their special needs or due to safety reasons. It is therefore important to know what the treatment needs are of these youths, and how their social environment may be supported and strengthened in order to prevent residential out-of-home placement, for instance by applying formal (Raposa et al., in press) or informal (Van Dam et al., 2017; 2018) mentoring. It must also become clear under what conditions youths at risk for residential placement cannot be treated through (forensic) foster care or home-based interventions, including family-style group care (Leloux-Opmeer et al., 2017). And if so, in what way (foster) families can be supported to overcome the risks of placement breakdown (See Konijn et al., 2019). Notably, the views and experiences of the youth themselves and their parents cannot be ignored when developing the most appropriate care for each youth. Lastly, Whittaker and colleagues (2016) state, that if treatment within residential care is unavoidable, it is useful to (1) offer help in closer collaboration with parents and other informal social network members, while the safety of the youth remains guaranteed, (2) make sure (therapeutic) residential care meets quality standards, is carefully monitored and properly designed, and (3) add intensive (foster) family-based interventions.

4.2. Limitations

The present meta-analysis has a number of limitations that need to be discussed in order to be able to fully appreciate our meta-analytic results, and prevent overinterpretation of our research findings. Notably, several limitations are shortcomings of the primary studies included in our meta-analytic review. Unfortunately, we could not include characteristics of residential care (e.g., level of security, availability of evidence-based treatment, the distinction between large scale institutional youth care and small-scale therapeutic residential care, group climate), intelligence of the youth, treatment integrity and length of residential stay as moderators in our analyses, because the included articles did not report sufficient data on these characteristics. We are
aware that the content of residential youth care may vary between different organizations, and that residential care is not as well researched as TFCO-A and the different forms of home-based treatment, and is rarely manualized, which may explain differences in outcomes of residential and non-residential care, such as better results for youth receiving TFCO-A than residential youth care (the present meta-analysis). However, ‘established’ and standardized non-residential interventions are often not carried out with high levels of treatment integrity, rendering these interventions ineffective, in particular for youth with conduct problems (see e.g., Goense et al., 2016). Moreover, Weisz et al. (2017) conducted a comprehensive meta-analysis of five decades of research on protocolled youth psychological therapy, showing no positive outcomes for youth with complex problems, in fact, those children or adolescents who may be at risk for out-of-home placement, and receive residential care, foster care or home-based care. Nevertheless, a meta-analysis by Van Stam et al. (2014) on the effectiveness of EQUIP and a study by Hoogsteder, Stams, Schippers and Bonnes (2018) on the effectiveness of Responsive Aggression Regulation Therapy showed positive effects of established manualized residential treatment on criminal recidivism in detained juvenile offenders. De Swart et al. (2012) made an attempt to compare evidence-based residential treatment with evidence-based non-residential treatment, but they found only one study (Wilmshurst, 2002), favoring home-based treatment at one year follow-up: youth receiving home-based treatment showed a reduction in symptoms of ADHD and internalizing problems, whereas youth receiving the residential program showed a clinical deterioration, resulting in increased internalizing problems. However, the goal of the present meta-analysis was to compare youth care in a professional residential setting, regardless of the specific content, with intensive home-based treatment (IHBT) and Treatment Foster Care Oregon for Adolescents (TFCO-A).

We solely included published studies in our meta-analysis, reducing the likelihood that results were based on lower quality research that had not been peer-reviewed, yet increasing the possibility of inherent publication bias. However, we found no indication of publication bias. Furthermore, although initial differences in problems at admission of youth served through non-residential and residential youth care were controlled through matching procedures, randomization and control for pre-test scores on outcome variables, we could not examine through moderator analysis whether the seriousness of the problems at admission affected the treatment outcome because a differentiation in relatively mild, moderate or high problem severity was not possible given the information provided in the included articles. Since, for example, Van der Pol and colleagues (2017) found that adolescents with more severe problems benefited more from MDFT, we recommend a fine-grained assessment of problem severity to be included in future studies on residential and non-residential care as a possible moderator of treatment outcome.

We were unable to conduct moderator analyses to distinguish between the effects of IHBT, MDFT, FFT, and MST due to a lack of studies, and the need for sufficient statistical power. However, findings of Van der Pol and colleagues (2019) show that these home-based interventions have much in common. They substantial overlap in the mechanisms and techniques used in MST, FFT, MDFT, brief strategic family therapy (BSFT), and even TFCO-A. For example, the mechanisms engagement, alliance, and interactional focus, and the techniques conflict management and communication skills were identified in all five treatments. Furthermore, there were twelve techniques and mechanisms found in four out of five treatment manuals, which further demonstrates the strong overlap between these interventions. Moreover, Van der Pol and others (2017) showed that MDFT and other multiple systems-based treatment, such as MST (Van der Stouwe et al., 2014), all have similar small effects on substance abuse, family functioning, internalizing and externalizing behavioral problems, and delinquency.

Finally, not all youth at risk for residential placement may have been included in the studies that are reviewed in this meta-analysis, since they could not participate in a randomized control trial for ethical reasons or because judicial measures prevented participation. Moreover, matched-control studies only compare a specific part of the population of youth at risk for residential placement, probably not the most and least severe cases. These drawbacks limit the generalizability of our study findings, and indicate that there might still be adolescents with complex needs who should be referred to residential youth care facilities or sentenced to detention.

4.3. Conclusion

Our findings suggest that youth with complex problems should be helped through non-residential care, especially treatment foster care, because this care yields slightly better results than residential care in reducing different types of behavioral problems, without the possible iatrogenic effects of residential care. These findings suggest that a part of residential care can be replaced by TFCO-A. Only in exceptional cases, when (treatment) foster care or home-based care cannot meet the safety needs of the youth, or is prohibited by judicial measures, residential youth care may be considered (Ainsworth, 2017). It is of major importance that alternative interventions be further developed and evaluated for those youth who are seen as unsuitable for foster care and home-based care. This is in line with our findings of better treatment outcomes for non-residential care, where (foster) family-based interventions are used, in close collaboration with parents and the informal network.

Funding

This study was funded by the Reformed Civil Orphanage Rotterdam, the Netherlands.

CRediT authorship contribution statement

Raymond V. Gutterswijk: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Project administration, Funding acquisition. Chris H.Z. Kuiper: Conceptualization, Methodology, Validation, Writing - review & editing. Navisha Lautan: Formal analysis, Investigation, Data curation. Elsemieke G. Kunst: Formal analysis, Investigation, Data curation. Frank C.P. van der Horst: Conceptualization, Methodology, Validation, Writing - review & editing. Geert Jan J.M. Stams: Conceptualization, Methodology, Validation, Formal analysis, Data curation, Writing - review & editing. Peter Prinzie: Conceptualization, Methodology, Validation, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Appendix A

Study characteristics of the studies in the meta-analysis

| Authors and year of publication | World part of origin | Comparison | Design | Average age (years) | Gender | N of effect sizes |
|---------------------------------|----------------------|------------|--------|---------------------|--------|-----------------|
| Barth et al. (2007)              | North America        | HBT vs group care | Matched control | 0-16 | mixed | 1 |
| Bergstrom & Höjman (2016)        | Western Europe       | TFCO-A vs group care | RCT | 12-17 | mixed | 2 |
| Cameron, Frensch, Preyde, & Smit Quosai (2011) | North America | HBT vs group care | Non-matched | 11.55 | mixed | 3 |
| Chamberlain & Reid (1998)        | North America        | TFCO-A vs group care | RCT | 11.4 | female | 4 |
| Chamberlain, Leve, & De Garmo (2007) | North America | TFCO-A vs group care | RCT | 15.3 | female | 4 |
| Eddy, Whaley, & Chamberlain (2004) | North America | TFCO-A vs group care | RCT | 14.9 | male | 3 |
| Harold et al. (2013)             | North America        | TFCO-A vs group care | RCT | 12.9 | mixed | 6 |
| Henggeler et al. (1999)          | North America        | HBT vs group care | RCT | 12.9 | mixed | 6 |
| Henggeler et al. (2002)          | North America        | HBT vs group care | Matched control | 8.10 | mixed | 5 |
| Henggeler, Milton, & Smith (1992) | North America | HBT vs group care | RCT | 15.2 | mixed | 4 |
| Leve, Chamberlain, & Reid (2005) | North America        | TFCO-A vs group care | RCT | 15.3 | female | 3 |
| Liddle et al. (2018)             | North America        | HBT vs group care | RCT | 15.4 | mixed | 20 |
| Mattejat, Hirt, Wilken, Schmidt, & Renschmidt (2001) | Western Europe | HBT vs group care | RCT | 11.9 | mixed | 4 |
| McCrave, Lee, Barth, & Ruskits (2010) | North America | TFCO-A vs group care | RCT | 11.1 | mixed | 5 |
| Portwood et al. (2018)           | North America        | TFCO-A vs group care | Non-matched | 13.81 | mixed | 8 |
| Poultan et al. (2014)            | North America        | TFCO-A vs group care | RCT | 15.31 | mixed | 8 |
| Preyde et al. (2011a)            | North America        | HBT vs group care | Matched control | 13.9 | mixed | 20 |
| Preyde et al. (2011b)            | North America        | HBT vs group care | Non-matched | 13.9 | mixed | 8 |
| Preyde, Adams, Cameron, & French (2009) | North America | HBT vs group care | Non-matched | 11.57 | mixed | 13 |
| Robet, Armstrong, & Dollard (2011) | North America | TFCO-A vs group care | Matched control | - | mixed | 2 |
| Robet, Armstrong, Dollard, & Rohrer (2013) | North America | TFCO-A vs group care | Matched control | 13.08 | mixed | 6 |
| Ryan, Marshall, Herz, & Hernandez (2007) | North America | HBT vs group care | Matched control | 7-16 | mixed | 1 |
| Wilmshurst (2002)                | North America        | HBT vs group care | RCT | 10.67 | mixed | 8 |

Appendix B. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.childyouth.2020.104950.

References

References marked with an asterisk (*) are included in the meta-analysis.

Ainsworth, F. (2017). For the few not the many: An Australian perspective on the use of therapeutic residential care for children and young people. *Residential Treatment for Children & Youth, 34,* 325–338. https://doi.org/10.1080/08865751.2017.1383568.

Armijo-Olivo, S., Stiles, C. R., Hagen, N. A., Biondo, P. D., & Cummings, G. G. (2012). Assessment of study quality for systematic reviews: A comparison of the Cochrane collaboration risk of bias tool and the effective public health practice project quality asessment tool. *Journal of Evaluation in Clinical Practice, 18,* 12–18. https://doi.org/10.1111/j.1469-7610.2010.01516.x.

Asscher, J. J., Van Vugt, E. S., Stams, G. J. J. M., Deković, M., Eichelsheim, V. I., & Yousfi, S. (2015). Preventie van gesloten plaatsing van jongeren met ernstige gedragsproblemen in Friesland: Een pleidooi voor een gerichter aanpak [Prevention of residential placement of adolescents with severe behavior problems Friesland, a province of The Netherlands: A plea for a targeted approach]. *Orthopedagogiek: Onderzoek en Praktijk,* 54, 318–330.

Bruning, M. R., & Jong-de Krijn, M. P. (2015). From youth services to youth help: The new arrangement considered further. *Journal of Family- and Juvenile Law, 33,* 134–143.

Busschers, I., & Konijn, C. (2019). Inzet van spoedplaatsingen in JeugdzorgPlus: Het hoe, wat en waarom van het aanvragen van een gesloten spoedplaatsing voor een jongere door gezinsmanagers van Jeugdbescherming Regio Amsterdam [Emergency placement in residential youth care: The how, what and why of emergency residential care placements of adolescents by family managers of the Youth Protection Agency Amsterdam] (Report). Retrieved from https://www.projeventuste-amsterdam.nl/fileupload SpoedplaatsingeninJeugdzorgPlus_ JeugdbeschermingenDeKoppeling_2019.pdf.

Cameron, Frensch, K., Preyde, M., & Smit Quosai, T. (2011). Community adaptation of youth accessing residential programs or a home-based alternative: Contact with the law and delinquent activities. *Residential Treatment for Children & Youth,* 28, 150–175. https://doi.org/10.1080/08865751.2011.577744.

Chamberlain, P., & Reid, J. B. (1998). Comparison of two community alternatives to institutional youth care: The how, what and why of emergency residential care placements of adolescents by family managers of the Youth Protection Agency Amsterdam. *British Journal of Social Work,* 28, 1–13. https://doi.org/10.1093/bjsw/28.1.1.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and violent recidivism: A meta-analysis. *American Journal of Orthopsychiatry,* 69, 455–463. https://doi.org/10.1111/j.1939-0025.1999.tb01271.x.
