The outcome of institutional youth care compared to non-institutional youth care for children of primary school age and early adolescence

A multi-level meta-analysis


Published in:
Children and Youth Services Review

DOI:
10.1016/j.childyouth.2015.09.018

Link to publication

Citation for published version (APA):
The outcome of institutional youth care compared to non-institutional youth care for children of primary school age and early adolescence: A multi-level meta-analysis

E.L.L. Strijbosch a,*, J.A.M. Huijs b, G.J.J.M. Stams b, I.B. Wissink b, G.H.P. van der Helm c, J.J.W. de Swart d, Z. van der Veen b

a Juzt (Quality and Innovation unit), Erasmusweg 34, 4834 AA Breda, The Netherlands
b University of Amsterdam (Department of Forensic Child and Youth Care Sciences), Nieuwe Achtergracht 127, 1018 WS Amsterdam, The Netherlands
c Leiden University of Applied Sciences (Youth Expert Centre), Zernikedreef 11, 2333 CK Leiden, The Netherlands
d Windesheim University of Applied Sciences, Campus 2-6, 8017 CA Zwolle, The Netherlands

1. Introduction

Under the United Nations Convention on the Rights of the Child, children have the right to grow up in a stable and safe environment where they receive the warmth and support they need for their development (Children’s Rights Alliance, 2010; Höfte, Van der Helm, & Stams, 2012; United Nations, 1989). Unfortunately, not every parent is able to offer a stable and secure home, and some children have to live in foster care or institutional youth care (Manso, García-Baamonde, Alonso, & Barona, 2011).

There is an ongoing debate about the appropriateness of institutional care for children, (Chance, Dickson, Bennett, & Stone, 2010; Dozier et al., 2014; Souverein, Van der Helm, & Stams, 2013). Mainly since the last decennium, but beginning in the 1980s, there has been a shift from institutional towards community based care, and a change from a deficit-focused to a strength-focused approach, in particular building on family strengths and resources (Knappe, Kumpfer, & Alvarado, 2003; Leichtman, 2008; Lonne, Parton, Thomson, & Harries, 2009; Lösel & Farrington, 2012; Melkman, 2015; Weick, Rapp, Sullivan, & Kisthardt, 1989). From this perspective, a growing number of (evidence-based) treatment alternatives have been developed, such as Multi-Systemic Therapy (MST; Henggeler, Pickrel & Brondino, 1999b; Van der Stouwe, Asscher, Stams, Dekovic, & Van der Laan, 2014) and Multidimensional Treatment Foster Care (MTFC; Chamberlain & Reid, 1998). However, the appropriateness of institutional youth care compared to non-institutional youth care should still be judged taking the type and severity of the problems of children into account.

Compared to children in non-institutional care, children in institutional care show more aggressive behaviour, and have more often been diagnosed with oppositional defiant disorder or conduct disorder (Handwerk, Field, & Friman, 2001; Lee & Thompson, 2007; Vermaes & Nijhof, 2014). Recent studies show that the severe behaviour problems can be associated with abnormal brain development as a result of neglect and traumatization (Fairchild et al., 2013; Raine, 2013). Providing the right treatment for children in institutional care is therefore very complex. Besides, living in an institutional setting can in itself have a negative or positive impact on the development of children (Dunn, Culhane, & Taussig, 2010; Preyde, Adams, Cameron, & Frensch, 2009).

As a result of the separation from their parents, children may develop internalizing problems (White & King, 2011), externalizing problems (Van der Helm, Stams, & Van der Laan, 2011) and attachment problems (Johnson, Browne, & Hamilton-Giachritsis, 2006; Van den Bergh, Weterings, & Schoenmakers, 2011; Van den Dries, Juffer, Van Ijzendoorn, & Bakermans-Kranenburg, 2009). Also, negative peer influences, such as ‘deviancy training’ (Dishion, McCord, & Poulin, 1999), can affect the development of children in institutional care. Children’s aggressive behaviour can trigger coercive behaviour in professionals, with a detrimental effect on the living group climate. The institutional setting can, on the other hand, also provide the safety and protection children coming from harmful circumstances need. For a discussion on the negative and positive consequences of institutional youth care, see Souverein et al. (2013).

There is little consensus in the literature about the effectiveness and appropriateness of institutional youth care compared to non-institutional youth care, and how the above-mentioned problems are being addressed (Preyde et al., 2011; Souverein et al., 2013). In particular, the long-term outcomes for children and adolescents living in institutional youth care have been questioned (Chor, McClelland, Weiner, Jordan, & Lyons, 2012; Dregan & Gulliford, 2012; Frensch & Cameron, 2002). There are some studies that indicate positive outcomes, but they are mostly based on small samples, and control groups are often missing (Bean, White, & Lake, 2005). Some pre-experimental studies showed a reduction of behavioural and emotional problems after treatment in institutional youth care (Larzelere, Daly, Davis, Chmelka, & Handwerk, 2004; Leichtman, Leichtman, Barber, & Neese, 2001). More recently, Dregan, Brown and Armstrong (2011) have investigated the effectiveness of institutional youth care and foster care, and showed that children in both conditions were at increased risk of behavioural and emotional problems in adulthood. Relatively better outcomes were related to the involvement of families during placement, e.g., by offering family therapy (Chance et al., 2010; Schubert, Mulvey, Loughran, & Losoya, 2012). Also a short length of stay, a positive living group climate, aftercare services and minimizing placement instability were important factors associated with better outcomes in institutional youth care (Hoagwood & Cunningham, 1993; Kho et al., 2012; Schubert et al., 2012).

Many studies are less positive about institutional youth care. Negative peer influences are often mentioned (Aguilar-Vafaie, Roshani, Hassanabadi, Masoudian, & Afruz, 2011; Orobio de Castro, Merk, Koops, Veerman, & Bosch, 2005; Whitehead, Keshet, Lombowski, Domenico, & Green, 2007). Whitehead et al. (2007) maintained that institutional youth care focuses too much on the child itself instead of on the entire child system (peers, school, and parents). Additionally, Manso et al. (2011) showed that many children in institutional care do not only have problems with their personal and social functioning, but also have educational problems. Dregan and Gulliford (2012) concluded that children in institutional care develop less favourably compared to children in foster care. As a possible explanation for this result, they mentioned that foster care provides more positive care experiences because it is a relatively stable placement with early admission to care and, as opposed to institutional care, a limited number of different caregivers. Their study did not adjust for pre-care characteristics. As some studies indicate (Barth,
Greeson, Green, Hurley, & Sisson, 2007; Berger, Bruch, Johnson, James, & Rubin, 2009; Vermaes & Nijhof, 2014), children who are most disturbed and in need of specialized treatment are often not selected for foster care.

1.2. Institutional care for children: EBT versus CAU

Even though research findings on the effectiveness of institutional versus non-institutional care are not consistent, it can be assumed that institutional care will be needed for a certain group of children, and it is therefore important to address the differences between several kinds of institutional care. Institutional youth care is carried out in 24-hour care facilities for children and youth with emotional and behavioural problems (Boendermaker, Van Rooijen, & Berg, 2012; Lee & Thompson, 2007). These facilities aim to provide a safe and structured environment in which the child receives either regular (long-term) care to be able to grow up in a sound manner (institutional CAU), or regular care combined with (short-term) specific child- and family-centred treatments from multidisciplinary teams that create individual care or treatment plans for each child (Preyde et al., 2009), that is, institutional EBT. Institutional EBT contains group as well as individual treatment, which is mostly based on behavioural, cognitive and solution focused treatment models (Rose, 2014; Van der Helm & Hanrath, 2011; Whitlaker, Del Valle, & Holmes, 2015). For a recent overview of evidence-based treatment methods applied in institutional settings, see James, Alemi, and Zepeda (2013). Examples of group based programs are EQUIP (Van Stam et al., 2014) and Re-Art (Hoogstedet et al., 2014). In many economically developed countries the services provided by the institutions are defined, regulated and (governmentally) monitored on the basis of guidelines, principles and standardized assessment of performances (Behar, Friedman, Pinto, Katz-Leavy, & Jones, 2007; Godbrandsson, 2008).

In their meta-analysis on the effectiveness and implementation of evidence-based practices in institutional care settings, James et al. (2013) demonstrated that EBT can be implemented and tested within the context of institutional care, and notwithstanding the limitations within the underlying studies, they found overall encouraging outcomes that indicated improvements in multiple domains of functioning. However, this meta-analysis mainly concerned (late) adolescents, and may not be representative for children of primary school age and early adolescence.

1.3. Selection factors

As mentioned earlier, the type and severity of the problems, as well as the age of the child may determine whether a child is referred to institutional or non-institutional care. Youths who are placed out of their family home tend to display more problems, fewer strengths, and more risk factors than youth who remain at home (Farmer, Mustillo, Burns, & Holden, 2008). Risk factors may pertain to child maltreatment, persistent juvenile delinquency or both. Notably, Asscher, Van der Put, and Stams (in press) showed that many juvenile delinquents, in particular girls, have a history of child maltreatment. Stith et al. (2009), in their meta-analytic review, found that parent characteristics (parent anger/hyper-reactivity, parent perceiving the child as a problem, parent’s level of stress, parent self-esteem) and family factors (high family conflict, low family cohesion, low parent–child relationship quality) were most predictive of physical abuse and neglect. Assink et al. (2015) conducted a meta-analytic review on risk factors for persistent juvenile delinquency, and they found relatively large effects for the criminal history, aggressive behaviour, and alcohol/drug abuse domains, and small effects for the family, neurocognitive and attitude domains. The physical health, background and neighbourhood domains did not yield significant effects. Farmer, Southerland, Mustillo, and Burns (2005) showed that the likelihood of moving back home after institutional care was related to race, age, gender, family income and total child behaviour problems. Decreased strengths and more child-level risk factors predicted instability of reunification. Generally, the accumulation of risk factors may result in less positive outcomes (Evans, Li, & Whipple, 2013). To summarize, parent and family factors as well as behavioural problems, skills and delinquency may explain referral to institutional or non-institutional care. The extent to which the (dynamic) risk factors are diminished during treatment, as well as certain (static) background variables, may influence the chance of moving back home successfully.

1.4. Previous meta-analyses addressing institutional versus non-institutional care

Although there is a vast body of research examining the outcomes of institutional care for children and adolescents, it is still unclear whether institutional youth care yields better or worse outcomes compared to non-institutional youth care (Preyde et al., 2009). Institutional youth care is very costly (Frensch & Cameron, 2002), and with a shift towards more community and strengths based care, it is important to investigate outcomes of institutional youth care compared to alternatives, by means of meta-analytic reviews (James, 2011; Ziviani, Feeney, Cuskelley, Meredith, & Hunt, 2012).

Recently two meta-analytic studies have been performed comparing institutional youth care to non-institutional youth care. Van den Dries et al. (2009) focused on children under the age of four and found that, compared to children raised in institutions, (early) adopted children had more secure attachment relationships with their caregivers. De Swart et al. (2012) examined outcomes of institutional youth care over the past three decades for youth and young adults up to 23 years of age. In this study, it was concluded that institutional youth care can be equally effective as non-institutional youth care. Furthermore, De Swart et al. (2012) showed that institutional evidence-based treatment yielded better outcomes than institutional care as usual.

The study of De Swart et al. (2012) focused on a broad age range, and did not differentiate between several kinds of outcomes within the studies. For several reasons, which are outlined in the next paragraph, it is important to address outcomes of institutional care compared to non-institutional care for primary school age children and early adolescents in a separate meta-analytic study. This will add to existing knowledge about the effects of institutional youth care, and hopefully assist those who refer children to institutional care or other types of care.

1.5. The present study

The present study examines the outcomes of institutional youth care compared to non-institutional care and the outcomes of institutional evidence-based treatment (EBT) compared to institutional care as usual (CAU) for primary school age children and early adolescents. This meta-analysis is based on the meta-analysis of De Swart et al. (2012), but there are two main differences. First, the present meta-analysis focuses on the younger age group (children mainly in the range from 4–14 years, with a maximum of 17 years), whereas De Swart et al. (2012) primarily focused on adolescents and young adults (mainly in the range from 14–18 years, with a maximum age of 23). These different age groups have different cognitive abilities and developmental tasks and are subject to different socialization practices. Children and young adolescents are thought to be more vulnerable to the separation from their parents, whereas (late) adolescents regularly have formed their own identity and have attained more independence from their parents (De Wit, Slot, & van Aken, 2013). Because of these specific differences, it is important to increase the knowledge on effectiveness of institutional care in the younger age group, and be able to distinguish between different outcome measures. The second main difference from De Swart et al. (2012) is that the present study uses multilevel techniques in order to be able to include more effect sizes of the same study, increasing statistical power and allowing the
examination of more moderators than can be achieved in a standard meta-analysis.

The purpose of this multilevel meta-analysis is to examine the effects of institutional youth care on behaviour problems (externalizing, internalizing and total), skills (social and cognitive) and delinquency in primary school age children and early adolescents. To accomplish this, four comparisons have been made. First, institutional Evidence-Based Treatment (EBT), i.e., structured (individual and group) treatment based on theoretical and empirical evidence (mostly behavioural, cognitive and solution focused therapies), is compared to institutional Care As Usual (CAU), i.e., group care offering daily care and structure, mostly in a (psychiatric) living group setting. Second, institutional EBT is compared to non-institutional EBT, such as Functional Family Therapy. Third, institutional CAU is compared to non-institutional CAU, mostly foster care. In addition, the following moderators were also examined because of their possible relevance for the interpretation of the outcomes: year of publication, journal impact factor, study quality, study design, time of measurement, type of intervention, data source, sex, mean age, percentage of girls, target group, ethnicity and control for pre-test differences in outcomes between the experimental and control group.

2. Methods

2.1. Study selection

Three search methods were used to collect the studies. First, we collected the studies with children and youth between 4 and 17 years old that De Swart et al. (2012) included in their meta-analysis. Second, we searched for studies on a broad domain of institutional youth care in the period from 1970 to 2013 in electronic databases: ScienceDirect, Psyclnfo, Picarta, Springerlink, ERIC, Medline and Google Scholar. We used the following keywords in various combinations: residential care, institutional youth care, institutional youth care, institutional youth care, institutional youth care, institutional youth care, institutional youth care, institutional youth care, institutional youth care, institutional youth care, institutional youth care. The search led to a total of 9 new studies that had not been included in the De Swart et al., 2012 meta-analysis. From the meta-analysis of De Swart et al., 15 studies were selected. All 24 studies were read by three researchers. In several meetings the researchers reached consensus on which of the studies would be included. Out of the 9 new studies, 4 could be included in the meta-analysis. Reasons for exclusion were mainly that the maximum age of the total population was too high, or that the study compared different groups but did not have a control group. This strategy finally resulted in a total of 19 studies (N = 15,526 children and youth) meeting the inclusion criteria (see Table 1).

2.2. Inclusion criteria

Studies were included if (1) the (quasi-) experimental group received institutional EBT or institutional CAU; (2) the control group received institutional non-EBT, or non-institutional CAU; (3) the average age of the children was under 15, and the total age range fell between 4 and 17 years; (4) studies provided at least post-test scores or follow-up scores in order to be able calculate effect sizes. Each study included in the meta-analysis was coded for methodological and sample characteristics. Methodological characteristics were year of publication, journal impact factor, study quality, study design, time of measurement, type of comparison, type of intervention, and the data source. A study quality index (Q) was used to evaluate the study quality (Downs & Black, 1998). Sample characteristics were mean age, sex, percentage of girls, target group, ethnicity and the outcome variables. These moderators were divided into continuous moderators and discrete moderators. Continuous moderators were year of publication, journal impact factor, study quality, mean age, percentage of girls and ethnicity (percentage of non-Caucasian participants). Discrete moderators were sex (boys or mixed), type of measurement (post-test or follow-up), control on pre-test (yes or no), type of comparison (institutional EBT vs non-institutional EBT), institutional EBT vs institutional CAU, institutional CAU vs non-institutional EBT, institutional CAU vs non-institutional CAU. Study design (matched, non-matched or RCT), type of intervention (behavioural treatment, cognitive behaviour therapy, skills training, system treatment or no treatment), target group (civil, criminal, psychiatric or mixed), data source (official report, parental report, professional report or mixed) and the outcomes (total behaviour problems, externalizing behaviour problems, internalizing behaviour problems, social skills, cognitive skills and delinquency).

2.4. Publication bias

Studies reporting strong significant associations are more likely to be accepted for publication in a journal. Therefore, studies that report less strong significant associations are more difficult to find. Consequently, conclusions of this meta-analysis may be incomplete, which is called the file drawer problem (Rosenthal & Hersen, 1979). File drawer bias was examined using a funnel plot of the distribution of effect sizes. Each individual study’s effect size is plotted on the vertical axis against its sample size, standard error or precision (the reciprocal of the standard error) on the vertical axis. The distribution of effect sizes should be shaped as a funnel if no publication bias is present since the more numerous studies with small sample sizes are expected to show a larger variation in the magnitude of effect sizes than the
Table 1
Study characteristics.

<table>
<thead>
<tr>
<th>Study (publication year)</th>
<th>Impact factor</th>
<th>Study quality rating</th>
<th>World part of origin</th>
<th>Comparison</th>
<th>Design</th>
<th>Time of measurement</th>
<th>Intervention</th>
<th>N</th>
<th>Sexes</th>
<th>Average age</th>
<th>Percentage non-Caucasian</th>
<th>Target group</th>
<th>Control on pre-test</th>
<th>N of effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barth et al. (2007)</td>
<td>2.21</td>
<td>14</td>
<td>North America</td>
<td>Inst CAU – Noninst EBT</td>
<td>Matched control</td>
<td>Follow-up</td>
<td>Behaviour modification</td>
<td>786</td>
<td>Mixed</td>
<td>11.42</td>
<td>39.35</td>
<td>Civil</td>
<td>No</td>
<td>1</td>
</tr>
<tr>
<td>Berger et al. (2009)</td>
<td>3.63</td>
<td>16</td>
<td>Western Europe</td>
<td>Inst CAU – Noninst CAU</td>
<td>Matched control</td>
<td>Follow-up</td>
<td>Behaviour modification</td>
<td>2453</td>
<td>Mixed</td>
<td>9.62</td>
<td>52</td>
<td>Civil</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>French, Cameron, and Preyde (2009)</td>
<td>0.00</td>
<td>15</td>
<td>North America</td>
<td>Inst EBT – Noninst EBT</td>
<td>Non-matched</td>
<td>Post-test and follow-up</td>
<td>CBT</td>
<td>210</td>
<td>Mixed</td>
<td>11.55</td>
<td>2</td>
<td>Civil</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td>Henggeler et al. (1999a)</td>
<td>4.98</td>
<td>13</td>
<td>North America</td>
<td>Inst CAU – Noninst EBT</td>
<td>RCT</td>
<td>Post-test</td>
<td>Behaviour modification</td>
<td>113</td>
<td>Mixed</td>
<td>13.00</td>
<td>66</td>
<td>Psychiatric</td>
<td>Yes</td>
<td>5</td>
</tr>
<tr>
<td>James, Roosch, and Zhang (2012)</td>
<td>1.28</td>
<td>17</td>
<td>North America</td>
<td>Inst CAU – Noninst CAU</td>
<td>Matched control</td>
<td>Post-test</td>
<td>Behaviour modification</td>
<td>1191</td>
<td>Mixed</td>
<td>8.95</td>
<td>58.70</td>
<td>Civil</td>
<td>No</td>
<td>3</td>
</tr>
<tr>
<td>Kolko, Loar, and Sturkik (1990)</td>
<td>4.98</td>
<td>16</td>
<td>Western Europe</td>
<td>Inst EBT – Inst CAU</td>
<td>Matched control</td>
<td>Post-test</td>
<td>Behaviour modification</td>
<td>56</td>
<td>Mixed</td>
<td>10.35</td>
<td>50.32</td>
<td>Psychiatric</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Mattejat, Hirt, Wilken, Schmidt, and Remschmidt (2001)</td>
<td>2.38</td>
<td>15</td>
<td>Western Europe</td>
<td>Inst CAU – Noninst EBT</td>
<td>RCT</td>
<td>Post-test</td>
<td>Behaviour modification</td>
<td>68</td>
<td>Mixed</td>
<td>11.90</td>
<td>46</td>
<td>Unknown</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>McCrae, Bethany, Barth, and Rauktis (2010)</td>
<td>0.67</td>
<td>18</td>
<td>North America</td>
<td>Inst CAU – Noninst CAU</td>
<td>Matched control</td>
<td>Post-test</td>
<td>Behaviour modification</td>
<td>124</td>
<td>Mixed</td>
<td>12.25</td>
<td>67.70</td>
<td>Civil</td>
<td>Yes</td>
<td>5</td>
</tr>
<tr>
<td>Moody (1997)</td>
<td>0.00</td>
<td>11</td>
<td>North America</td>
<td>Inst EBT – Inst CAU</td>
<td>Non-matched</td>
<td>Post-test</td>
<td>Behaviour modification</td>
<td>28</td>
<td>Boys</td>
<td>14.30</td>
<td>89.29</td>
<td>Criminal</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Preyde et al. (2011)</td>
<td>0.00</td>
<td>12</td>
<td>North America</td>
<td>Inst CAU – Noninst EBT</td>
<td>Non-matched</td>
<td>Post-test and follow-up</td>
<td>Behaviour modification</td>
<td>112</td>
<td>Mixed</td>
<td>11.57</td>
<td>Unknown</td>
<td>Civil</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Robst, Armstrong, and Dollard (2011)</td>
<td>1.27</td>
<td>16</td>
<td>North America</td>
<td>Inst CAU – Noninst CAU</td>
<td>Matched control</td>
<td>Post-test</td>
<td>Behaviour modification</td>
<td>842</td>
<td>Mixed</td>
<td>Unknown</td>
<td>52.20</td>
<td>Civil</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Scholte and Van der Ploeg (2003)</td>
<td>0.00</td>
<td>13</td>
<td>Western Europe</td>
<td>Inst EBT – Noninst EBT</td>
<td>Non-matched</td>
<td>Post-test</td>
<td>Skills</td>
<td>105</td>
<td>Mixed</td>
<td>14.90</td>
<td>35</td>
<td>Mixed</td>
<td>No</td>
<td>3</td>
</tr>
<tr>
<td>Thompson et al. (1996)</td>
<td>1.12</td>
<td>15</td>
<td>North America</td>
<td>Inst EBT – Inst CAU</td>
<td>Non-matched</td>
<td>Post-test</td>
<td>Skills</td>
<td>587</td>
<td>Mixed</td>
<td>14.55</td>
<td>29.60</td>
<td>Mixed</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Wilmshurst (2002)</td>
<td>0.00</td>
<td>17</td>
<td>North America</td>
<td>Inst EBT – Noninst EBT</td>
<td>Non-matched</td>
<td>Post-test and follow-up</td>
<td>Skills</td>
<td>65</td>
<td>Mixed</td>
<td>10.74</td>
<td>5</td>
<td>Civil</td>
<td>Yes</td>
<td>8</td>
</tr>
</tbody>
</table>
less numerous studies with large effect sizes. A violation of funnel plot symmetry reflects publication bias, that is, a selective inclusion of studies showing positive or negative outcomes (Sutton, Duval, Tweedie, Abrams, & Jones, 2000). Funnel plot asymmetry was tested by regressing the standard normal deviate, defined as the effect size divided by its standard error, against the estimate’s precision (the inverse of the standard error) that largely depends on sample size (see Egger, Smith, Schneider, & Minder, 1997). If there is asymmetry, the regression line does not run through the origin and the intercept significantly deviates from zero.

2.5. Analysis

For each of the studies Cohen’s d was calculated for the effectiveness of institutional youth care on the basis of differences between institutional and non-institutional youth care and difference between evidence-based institutional treatment and institutional care as usual. Both the post-test data and the follow-up data were used. Effect sizes were calculated on the basis of means and standard deviations, percentiles and t-, F-, χ²-, p-values. For this purpose, Wilson’s effect size determination program (Wilson, 2001) was used. An effect size of d = 0.20 is considered as small, an effect size of d = 0.50 is considered as medium and an effect size of d = 0.80 is considered as large (Cohen, 1988). Outliers were checked on the basis of z-values larger than 3.29 or smaller than −3.29 (p < 0.005; Tabachnick & Fidell, 2001). No outliers were found. After that, categorical variables were turned into dichotomous dummy codes and continuous moderator variables were centred around their mean in order to be able to conduct multilevel meta-analysis.

The homogeneity of the combined, total effect size was tested with a z-test of the between study variance (total study variance divided by its standard error). If this z-test is significant, there is heterogeneity. In case of significant heterogeneity, moderators may account for differences between studies, and it is imperative to conduct categorical and/or continuous moderator analyses.

Finally, multilevel analysis was conducted by using the program MLwiN (Hox, 2002). The multilevel random effects model takes the hierarchical structure of the data into account in which the effect sizes (the lowest level) are nested within studies (the highest level). Iterative maximum likelihood procedures were applied to estimate unknown parameters.

3. Results

This meta-analysis included 19 studies, comprising 63 effect sizes, with the results based on N = 15,526 children and youth between the ages of 4 and 17 years. Table 2 shows a representation of the overall mean effect size and the significant discrete moderators. The overall mean effect size was d = −0.02. The individual study effect sizes ranged from −1.14 to 1.56. Possible publication bias was examined by testing funnel plot asymmetry. The standard normal deviate was regressed against the estimate’s precision. As the intercept did not significantly deviate from zero (z = 1.879; p = 0.08), there was no indication of funnel plot asymmetry and therefore no indication of publication bias. These findings suggest that the mean effect size can be considered robust. Finally, the overall mean effect size proved to be heterogeneous (z = 2.875), indicating that the effect was not the same in all studies. This is a precondition to be able to expect significant outcomes from moderator analyses.

Table 2 shows that type of comparison was a significant moderator: χ²(2) = 8.932, p < .05. The reference category was the comparison between institutional evidence-based treatment and non-institutional evidence-based treatment (d = 0.342, ns). The institutional Care As Usual (CAU) versus non-institutional CAU comparison differed significantly from the reference group comparison (z = 2.601, p < .01), yielding a small-to-medium negative significant effect (d = −0.342; z = 2.280, p < .05). This means that children in non-institutional CAU, mostly foster care, had somewhat better outcomes than children in institutional CAU, whereas there were no significant differences in the effects of EBT between institutional and non-institutional care.

Another moderator variable with a significant effect was study design: χ²(2) = 9.656, p < .01. Non-matched studies differed significantly from matched studies: z = 3.217, p < .01. Matched studies yielded a negative and significant effect size (d = −0.309; z = 2.255), whereas non-matched studies showed a positive and significant effect size.

(d = 0.299; z = 2.30, p < .05). The non-matched studies showed better outcomes for children in institutional youth care and the matched studies showed better outcomes for children in non-institutional youth care.

Finally, type of outcome was a significant moderator: χ²(5) = 25.115, p < .001. Delinquency differed significantly from the reference

<table>
<thead>
<tr>
<th>Moderator variables</th>
<th># studies</th>
<th># ES</th>
<th>Mean d</th>
<th>z₁</th>
<th>RC</th>
<th>z₂</th>
<th>Heterogeneity</th>
<th>Fit χ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>19</td>
<td>63</td>
<td>−0.018</td>
<td>0.176</td>
<td></td>
<td></td>
<td>2.875**</td>
<td></td>
</tr>
<tr>
<td>Type of comparison</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.804**</td>
<td>8.932*</td>
</tr>
<tr>
<td>Inst EBT vs Noninst EBT (RG)</td>
<td>3</td>
<td>20</td>
<td>0.342</td>
<td>1.591</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inst EBT vs Inst CAU</td>
<td>4</td>
<td>8</td>
<td>0.285</td>
<td>1.397+</td>
<td>−0.057</td>
<td>0.193</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inst CAU vs Noninst EBT</td>
<td>6</td>
<td>22</td>
<td>−0.034</td>
<td>0.252</td>
<td>−0.310</td>
<td>0.166</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inst CAU vs Inst CAU</td>
<td>6</td>
<td>14</td>
<td>−0.342</td>
<td>2.280*</td>
<td>−0.675</td>
<td>2.601**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.814**</td>
<td>9.656**</td>
</tr>
<tr>
<td>Matched (RG)</td>
<td>7</td>
<td>15</td>
<td>−0.309</td>
<td>2.255*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-matched</td>
<td>8</td>
<td>35</td>
<td>0.299</td>
<td>2.300*</td>
<td>0.608</td>
<td>3.217**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td>4</td>
<td>13</td>
<td>−0.131</td>
<td>0.712</td>
<td>0.177</td>
<td>0.773</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of outcome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.833**</td>
<td>25.115***</td>
</tr>
<tr>
<td>Total problems (RG)</td>
<td>11</td>
<td>14</td>
<td>0.035</td>
<td>0.343</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Externalizing behaviour</td>
<td>10</td>
<td>13</td>
<td>0.168</td>
<td>1.680</td>
<td>0.133</td>
<td>2.180*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internalizing behaviour</td>
<td>10</td>
<td>13</td>
<td>−0.051</td>
<td>0.510</td>
<td>−0.085</td>
<td>1.393</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social skills</td>
<td>4</td>
<td>6</td>
<td>0.089</td>
<td>0.005</td>
<td>0.054</td>
<td>0.409</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive skills</td>
<td>9</td>
<td>12</td>
<td>0.051</td>
<td>0.400</td>
<td>0.016</td>
<td>0.208</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delinquency</td>
<td>5</td>
<td>5</td>
<td>−0.329</td>
<td>2.179*</td>
<td>−0.364</td>
<td>2.318*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. RG = reference group; # studies = number of studies; # ES = number of effect sizes; Mean d = mean effect size; z₁ = significance of moderator; RC = slope; z₂ = differences in mean d with reference group; Heterogeneity = within class heterogeneity (z); Fit = difference with model without moderators (χ²); + = one-sided trend = 0.08.

⁎ p < 0.05.
⁎⁎ p < 0.01.
⁎⁎⁎ p < 0.001.
group, which was total problems: $z = 2.318, p = 0.05$, yielding a small-to-medium negative significant effect ($d = -0.329; z = 2.179, p < 0.05$). This means that children in institutional youth care showed more delinquent behaviour compared to children in non-institutional youth care. The other outcome categories did not show a significant effect, and were therefore not mentioned in Table 2.

Table 3 shows the results for the continuous moderators. Year of publication was a significant moderator ($z = 2.692, p < 0.01$), indicating that earlier published studies were associated with larger effect sizes ($\beta_1 = -0.035$). Also with regard to sex of the child a significant moderating effect was found ($z = 2.500, p < 0.05$). Studies with a high percentage of females were associated with smaller effect sizes ($\beta_1 = -0.015$).

4. Discussion

The purpose of the present meta-analysis was to examine the outcome of institutional youth care compared to non-institutional care and the outcome of institutional evidence-based treatment (EBT) compared to institutional care as usual (CAU) for primary school age children and early adolescents. This differs from the meta-analysis of De Swart et al. (2012), which had a broader age range and focused primarily on middle and late adolescents and young adults. In addition, the present study uses multilevel techniques in order to be able to include more effect sizes of the same study, preserving all relevant information, which increases statistical power and enables the examination of more moderators than can be achieved in a standard meta-analysis.

The main conclusion is that institutional CAU showed a small-to-medium negative significant effect compared to non-institutional CAU ($d = -0.342$). This indicates that children in non-institutional CAU, mostly foster care, had somewhat better outcomes than children in institutional CAU, which is regular group care offering daily care and structure in a (psychiatric) living group setting. This conclusion is in line with the retrospective study by Dregan and Gulliford (2012).

They found that children in institutional youth care develop less favourably compared to children in foster care, and institutional care is associated with increased risk of adult criminal convictions and depression. Our findings differ from those in the meta-analytic study of the De Swart et al. (2012), which focused on a broader age range, with an overrepresentation of late adolescents and young adults. Although De Swart et al. (2012) found that non-institutional CAU had slightly better outcomes than institutional CAU, this difference just failed to reach statistical significance. Notably, selection effects should be taken into account when interpreting this result, because children with more complex problems generally require more intensive and specialized treatment (Evans et al., 2013; Farmer et al., 2008; Stith et al., 2009). This specialized treatment can be provided in non-institutional care, but may increase the risk for placement disruptions (Oosterman et al., 2007; Van Oijen, 2010; Vanschoonlandt, Vanderfaillie, Van Holen, De Maeyer, & Andries, 2012). As a result, these children may be referred less often to foster care. Sometimes children start in foster care and are later referred to institutional care because of the severity of their problems (Hussey & Guo, 2005).

No differences were found between institutional and non-institutional youth care when institutional treatment was evidence-based, which is in line with findings from the meta-analysis of De Swart et al. (2012). However, we did not find a significant advantage in providing youth with institutional EBT instead of institutional CAU. De Swart et al. (2012) found a significant difference between institutional EBT and institutional CAU for middle and late adolescents as well as young adults, with better outcomes for institutional EBT. In the present study the effect size was small-to-medium in favour of institutional EBT for children and young adolescents, but did not reach significance. A possible explanation for this result is a lack of treatment integrity, which is a shortcoming in the application of evidence-based treatment methods in institutional care (James et al., 2013). Another possible explanation is that group workers may not be responsive enough to the individual needs of the children in their group, which would be a basic requirement to be able to generate positive results of institutional EBT (Van der Helm, 2011).

Variations in the way children adapt to institutional care do not only depend on importation factors (Gover, MacKenzie, & Armstrong, 2000; Hussey & Guo, 2005), i.e. characteristics of individuals before entering institutional care shaping their adjustment, but also on environmental factors, such as the living group climate. Recent studies on this topic, mainly performed in forensic settings, suggest that group climate factors influence treatment outcomes. For example, Van der Helm, Klapwijk, Stams, and Van der Laan (2009) and Schubert et al. (2012) showed that a positive or ‘open’ living group climate characterized by warmth and responsiveness from group workers, opportunities for development, and a safe and structured environment, had a positive effect on the treatment of juvenile delinquents. A ‘closed’ or repressive climate, on the other hand, has been found to be associated with negative developmental outcomes and aggression (Van der Helm, Stams, Van Genabbeek, & Van der Laan, 2012). Therapeutic alliance has also been identified in the literature as a predictor of positive treatment outcomes and can be influenced (Harder, Knorth, & Kalverboer, 2012; Hurley, Lambert, Van Ryzin, Sullivan, & Stevens, 2013; Karver, Handelsman, Fields, & Bickman, 2006). In addition, although there is abundant evidence indicating the prevalence of trauma exposure among youth in institutional care, few models exist for creating trauma-informed milieu treatment (Brown, McCauley, Navalta, & Saxe, 2013), which could also hamper better results for institutional EBT.

Moderator analyses indicated that the design of the studies influenced the conclusion about the effectiveness of institutional youth care. Institutional youth care showed better outcomes in non-matched studies, whereas matched studies showed less favourable outcomes for institutional youth care. A possible explanation is that a relatively greater therapeutic gain can be achieved when problem rates are higher at the start, which is often the case in institutional care (Barth et al., 2007; Berger et al., 2009). The main benefit of matching studies is high internal conclusion validity, because the participants in the experimental group and control group are supposed to be comparable. For
this reason, more value is attached to the outcomes of the matched studies, which yield less positive results for institutional youth care. Notably, a possible drawback of matching is that children with the most severe problems are often left out of the analyses because they are less likely to match with juveniles in the control group. This may violate the external validity or, in other words, clinical representativeness of studies using a matched control design (Goodman et al., 1997).

In contrast to the study of De Swart et al. (2012), the present study showed a significant moderator effect for delinquency. Children in institutional youth care showed more delinquent behaviour compared to children in non-institutional youth care. Differences between institutional and non-institutional care with regard to delinquent problem behaviour may be more extreme in the younger age group, because presently a (young) child is only referred to institutional care when no other options are available. Other types of outcomes did not show significant differences between institutional and non-institutional care in this meta-analysis, which is in line with results reported by De Swart et al. (2012).

Finally, significant moderating effects were found for year of publication and sex of the child. Studies that were published more recently and studies with a higher percentage of females yielded smaller effect sizes. This first finding may be explained by the fact that statistical techniques, such as propensity score matching, have become more advanced over time, which enables better control for multiple confounders that are responsible for differences between institutional and non-institutional care. In addition, and probably even more important, growing insights into offering qualitatively good institutional care may have diminished the differences between institutional and non-institutional youth care. The association between sex of the child and magnitude of the effect sizes could be explained by the fact that girls in institutional care seem to suffer more from neglect (Gunther Moor, 2011) and more often have a history of trauma and sexual abuse, resulting in relatively more internalizing problems (Zurbriggen, Gobin, & Freyd, 2010). Research indicates that children with externalizing behavioural problems seem to make more progress in institutional care than youth with internalizing behaviours problems, which are mostly girls (Knorth, Harder, Zandberg, & Kendrick, 2007). This could be an explanation for the smaller differences between institutional and non-institutional youth care when the percentage of girls increases.

4.1. Implications for future research

In the first place, given the non-significant and small-to-medium effect size of the comparison between institutional EBT and institutional CAU, more research is necessary to examine the effectiveness of evidence-based treatment in institutional youth care for children. Future research should focus on the conditions that make established treatment methods work in institutional youth care for (young) children: what works for whom and under what circumstances? An interesting discussion on this topic, introducing a new model on translational research for the field of residential child care, is found in Nunno, Sellers, and Holden (2014).

Furthermore, in general, future effect studies focussing on institutional care versus alternatives should report more information on (cumulative) risk factors (Evans et al., 2013; Farmer et al., 2008; Sth et al., 2009), because changes in dynamic risk factors are critical to establish the effects of the care provided. In many effect studies hardly any information is reported on parent and family factors. Tools that assist clinicians in identifying and matching the level of risk to the intensity of treatment are available; see, for example Augimeri, Walsh, Woods, and Jiang (2012). In addition, protective factors should be examined in future studies, which is in line with the shift towards strengths-based care (Kumpfer & Alvarado, 2003). Results on changes in risk and protective factors would enable refinements of future meta-analytic studies.

Another recommendation for future effect studies is to use child self-report measures. Self-report measures can be a valuable supplement to the official, parental and professional reports (Dunn et al., 2010). Yet another important factor to be taken into account in future studies is IQ. Many children in institutional youth care have a mild intellectual disability, which can have a great impact on treatment results (Van Nieuwenhuijzen, Vriens, Scheepmaker, Smit, & Porton, 2011).

4.2. Implications for policy and practice

The main conclusion of this study is that non-institutional CAU had slightly better outcomes than institutional CAU, which supports the idea that institutional care should be considered as a last resort, to be used only when non-residential alternatives are less appropriate. When problems are too severe and institutional care seems to be the best choice given the situation, it is recommended that attention be paid to the living group climate, therapeutic alliances and trauma-sensitiveness as preconditions to be able to provide adequate (evidence-based) institutional care for children. These factors should be regularly monitored, and group workers should receive proper training and coaching in order to make necessary improvements and enable children to express their thoughts and feelings about living in a residential setting more easily (Hunt, 2010; Leichtman, 2008). For example, promoting positive outcomes for children with a special emphasis on developing safer environments and healthy relationships is described in the Children and Residential Experiences (CARE) programme (Holden, 2009).

4.3. Limitations of the study

A general limitation of meta-analysis is that it cannot account for all underlying differences between studies by means of moderator analyses, such as the impact of culture and care systems among countries (Weisz, Ugueto, Cheron, & Herren, 2013b). There are also some technical limitations of this meta-analysis that should be noted. First, in some moderator categories the number of effect sizes was very small, which may have resulted in less reliable outcomes. Second, not every study reported the range of the age of the children. When a mean age lower than 15 years was reported without the exact age range, the study was included. As a result, we do not know if the children included in these studies were mainly primary school age children or early adolescents, which hampers more specific generalizability of our study findings. Another limitation is that evidence-based treatment (EBT) and care as usual (CAU) are broadly defined categories. The limited number of studies and effect sizes did not allow making more refined comparisons between different kinds of treatments.

5. Conclusion

Despite its limitations, this meta-analysis provides new insights into the outcome of institutional youth care compared to non-institutional youth care and the outcome of institutional evidence-based treatment (EBT) compared to institutional care as usual (CAU) for children of primary school age and early adolescence. Also, the application of multilevel analysis made it possible to include more effect sizes than can be achieved with regular meta-analysis and to examine more moderators. This study indicates that a particular group of children seems to have better outcomes in non-institutional care as usual, such as foster care, compared to institutional care. An explanation is that a more stable (family) environment provides better conditions for development and treatment, such as the development of secure child–caregiver attachment relationships (Dozier et al., 2014). Additionally, within such an environment, children may be less affected by negative peer influences (Aguilar-Vafaie et al., 2011; Whitehead et al., 2007). When problems are too severe to treat at home or in foster care, specialized institutional treatment may still be required. In order to offer optimal (evidence-
based) care within institutions, one should besides child and family related characteristics (risk and protective factors) also take environmental factors (e.g., group climate) and therapeutic alliance into account, and provide a trauma-sensitive treatment milieu. Making careful decisions in referring children to specific forms of care prevents problems from getting worse later in life and diminishes recidivism and re-entry in care, thereby also diminishing costs for society in the long term.

Acknowledgements

We thank Martha J. Holden, MS, Extension Associate with the Bronfrenbrenner Center for Translational Research at the College of Human Ecology, Cornell University, for her help with reviewing the English in this article.

References


