

Accepted Manuscript

Cite as: Tarren-Sweeney, M., & Goemans, A. (2019). A narrative review of stability and change in the mental health of children who grow up in family-based out-of-home care. *Developmental Child Welfare*. 1(3) 273–294 DOI: 10.1177/2516103219874810

ABSTRACT

The present review sought to address the following questions: What evidence is there that long-term family-based out-of-home care (OOHC) has a general, population-wide effect on children’s mental health, such that it is generally reparative or generally harmful? Does entry into long-term OOHC affect children’s mental health, as evidenced by prospective changes over the first years in care? And, is the reparative potential of long-term, family-based OOHC moderated by children’s age at entry into care? Fourteen studies were identified for review. We found no consistent evidence that family-based OOHC exerts a general, population-wide effect on the mental health of children in care; or that entry into care has an initial effect on children’s mental health; or that children’s age at entry into care moderates their subsequent mental health trajectories. Instead, several longitudinal studies have found that sizeable proportions of children in care manifest meaningful improvement in their mental health over both short- and long-term timeframes; and that similarly sizeable proportions experience meaningful deterioration in their mental health. Rather than asking whether long-term family-based care is generally reparative or harmful for the development of previously maltreated children, future investigations should instead focus on identifying the systemic and interpersonal characteristics of care that promote and sustain children’s psychological development throughout childhood – and those characteristics that are developmentally harmful (i.e., for which children is the experience of care beneficial, and for which children is it not?). The review concludes with recommendations for the design of improved cohort studies that can address these questions.

Introduction

With increasing numbers of children growing up in long-term out-of-home care (OOHC), governments and children's agencies need better information about how the experience of growing up in care by children with prior exposure to severe social adversity affects their psychological development and well-being. The present article seeks to add to this information, by reviewing longitudinal studies that measured changes in children's mental health whilst residing in long-term, family-based (i.e. foster and kinship) OOHC. Estimating the reparative and harm potentials of long-term family-based OOHC requires an understanding of the relative long-term impact of pre-care and within-care experiences, including complex transactional mechanisms. The most critical developmental consequence of children's early exposure to severe social adversity is poor mental health. Numerous cross-sectional studies have established that children in care manifest high mean levels and rates of mental health difficulties. Though rates vary a little by survey and location, up to half of such children have clinical-level mental health difficulties, and another 20% to 25% have difficulties approaching clinical significance (Oswald, Heil, & Goldbeck, 2010).

In jurisdictions where children predominantly enter long-term care following severe and persistent maltreatment, a child's age at entry into care approximates their length of post-birth exposure to chronic and severe maltreatment. Furthermore, a child's 'age at entry into care' strongly predicts their subsequent mental health difficulties – with entry at younger age being protective (Burge, 2007; Hukkanen, Sourander, Bergroth, & Piha, 1999a; Tarren-Sweeney, 2008). This is consistent with cumulative trauma exposure models (Charlotte, Viding, Fearon, Glaser, & McCrory, 2017), neuroscience, and attachment theory. Regardless of prior conditions, the attachment systems of infants who enter foster care have been found to be responsive to changes in parenting style (Dozier, Stovall, Albus, & Bates, 2001). Conversely, a study of late-placed children found that the severity of their pre-care maltreatment was associated with their maternal and self-representations, which in turn predicted children's subsequent representations of their relationships with foster mothers, as well as their subsequent mental health (Milan & Pinderhughes, 2000). A range of psychological and neuro-biological processes in early childhood that are critical to human social functioning are impaired by early and prolonged exposure to traumatic maltreatment, and by the absence of nurturing, sensitive care. However, it is important to note that early exposure to severe and/or chronic maltreatment need not result in irreparable harm. Rather, there is emerging evidence that its effects manifest as *latent vulnerabilities* that are mitigated to varying degrees by children's subsequent experience of optimal developmental experiences (McCrory & Viding, 2015).

What then is known of the developmental effects of children's experiences in family-based OOHC? Attachment theory predicts that the developmental effects of OOHC should vary according to the characteristics of a child's attachment development prior to their entry into care,

notably their internal working model of attachment; and to caregiver sensitivity and their ability to provide a 'secure base' (Bowlby, 1988; Schofield, 2002). Many such children are *primed for insecurity* when they enter care, due to their compromised attachment development and distorted representations of caregivers and caregiving; as well as the loss of their parents and being placed with unfamiliar carers (Milan & Pinderhughes, 2000; van den Dries, Juffer, van IJzendoorn, & Bakermans-Kranenburg, 2009).

There is accumulating evidence that quality of caregiving, caregiver bonding, caregiver commitment, and maltreatment in care, are factors that directly influence children's felt security and psychological development, and regulate their potential to recover from attachment- and trauma-related psychopathology (Dozier, Grasso, Lindheim, & Lewis, 2007; Quiroga & Hamilton-Giachritsis, 2016; Tarren-Sweeney, 2008). Several longitudinal studies have also identified that children incur further deterioration in their mental health following placement disruptions, which is a common occurrence in OOHC (Aarons et al., 2010; Delfabbro & Barber, 2003; Newton, Litrownik, & Landsverk, 2000; Villodas, Litrownik, Newton, & Davis, 2016).

Rationale for the present review

A starting point for examining the developmentally reparative versus harmful effects of long-term family-based OOHC is prospective measurement of stability and change in children's mental health. In comparison to the large number of cross-sectional mental health surveys conducted with children in care, there have been relatively few prospective studies. Much of the available prospective data are compromised by high sample attrition, short prospective timeframes, and small sample size. A recent series of meta-analyses pooled prospective mean score changes in externalising difficulties (21 studies), internalising difficulties (24 studies), and total difficulties i.e. global mental health (25 studies) (Goemans, van Geel, & Vedder, 2015). These meta-analyses showed no statistically or clinically significant changes over time in children's internalizing (Hedges' $g = -.10$, 95% CI = $-.27; .07$, $p = .25$, $N = 1,984$), externalizing ($g = -.04$, 95% CI = $-.24; .15$, $p = .66$, $N = 1,729$), or total behaviour problems ($g = -.10$, 95% CI = $-.28; .07$; $p = .24$, $N = 2,523$). Various moderator analyses failed to show effects when comparing studies on study length, sample size, publication type, attrition, or mean age. Instead, the three meta-analyses identified considerable heterogeneity across the various study findings, with some reporting large mean increases in mental health scores over time, and others reporting large reductions (Goemans et al., 2015).

What might account for this? One explanation is that family-based OOHC does not exert a general, population-wide effect on children's development. Group mean score changes are really only informative if such changes reflect a general and largely uniform shift in the distribution of mental health scores over time i.e. if children's mental health *generally* improves or deteriorates whilst growing up in care. Secondly, discrepant findings may be accounted for by

variability in study design. A narrative review provides a vehicle for demystifying heterogeneous and discrepant findings using scholarly reasoning – that can yield additional insights to those afforded by meta-analyses.

The present review aims to address the following research questions with respect to children (including infants and adolescents), who are placed into long-term, family-based OOHC, following serious and/or chronic maltreatment.

1. What evidence is there that long-term family-based care has a general, population-wide effect on children's mental health, such that it is generally reparative or generally harmful?
2. Does entry into long-term OOHC affect children's mental health, as evidenced by prospective changes over the first years in care?
3. Is the reparative potential of long-term, family-based OOHC moderated by children's age at entry into care?

Review method

To perform the present review, we first conducted a literature search in PsychINFO, Medline and ERIC to identify studies published before July 2018, using the terms *longitudinal*, *prospective* or *repeated measures*, combined with the terms *out-of-home care*, *looked after children*, *foster care*, or *kinship care*. This was supplemented by manual searches of article references, and cross-checking studies included in the recent meta-analysis (Goemans et al., 2015). Subsequent checks were made to identify relevant studies published after July 2018, and prior to the present article being published.

Study inclusion/exclusion criteria

1. Sample predominantly experienced pre-care maltreatment. The review is focussed towards understanding how children who predominantly experience severe and/or chronic maltreatment in their parents' care *subsequently develop* in long-term family-based care. Longitudinal studies of the development of orphans or abandoned infants placed in foster care without previous history of maltreatment refer to a different population, and were excluded from the present review. Similarly, attempts were made to confirm that study samples were predominantly placed into care following maltreatment, either from information provided in the study publication, or from published descriptions of OOHC systems in the study locations. All of the studies located in the present literature search that met all other selection criteria were carried out in locations where children predominantly enter court-ordered care due to maltreatment.

2. Measures obtained whilst children resided in family-based OOHC. Studies were included in the present review if mental health data were measured on at least two occasions when children were residing in family-based OOHC. Four studies were excluded because some or

most of the study samples were residing in their birth parents' care at follow-up (Havnen, Breivik, & Jakobsen, 2014; Newton et al., 2000; Proctor, Skriner, Roesch, & Litrownik, 2010; Villodas et al., 2016).

3. Prospective measurement using the same type of informants. A further methodological problem is estimating prospective change from reports provided by *different types* of informants at different time points, such as comparing birth parent reports at baseline with foster carer reports at follow-up. Three studies were excluded because they estimated baseline mental health from parent-report scores, and follow-up mental health from foster carer-report scores (Berger, Bruch, Johnson, James, & Rubin, 2009; Havnen et al., 2014; Rubin, O'Reilly, Luan, & Localio, 2007).

4. Reliable informants of children's mental health. Two cohort studies estimated foster children's mental health from social worker reports. However, social workers do not have sufficient proximal engagement with children in care to reliably inform on their mental health (English & Graham, 2000; McCrae & Barth, 2008), and those studies are thus excluded from the present review (Barber & Delfabbro, 2005; Fanshel & Shinn, 1978; Frank, 1980). Similarly, it is doubtful that children's birth parents are able to accurately report on their children's mental health *when they are not* residing in their care, and one study that employed this method was excluded from the review (Linares, Li, ShROUT, Brody, & Pettit, 2007).

5. Sufficient sample size. Consideration was given to excluding studies with small samples that were insufficient for identifying effect sizes that are clinically meaningful. We calculated that the minimum sample size required for identifying a large effect size (defined as $d \geq 0.80$) with 95% confidence is $N=25$. Eight studies that retained very small samples ($N < 30$) at follow-up were excluded from the present review (Bogart, 1988; Gonzalez, 1999; Haight, Black, & Sheridan, 2010; Lawrence, Carlson, & Egeland, 2006; Leathers, Spielfogel, McMeel, & Atkins, 2011; Rubin et al., 2007; Rushton, Treseder, & Quinton, 1995; White, 1997).

6. Measures reported for common cohorts. Three studies reported baseline and follow-up data for non-identical groups, such that the differences in mean scores do not reflect average within-subject change (Fernandez, 2008; Hiller & St. Clair, 2018; Portwood et al., 2018).

7. Sufficient prospective timeframe. The shorter the prospective timeframe, the more likely it is that a study sample includes children in temporary and short-term care. Given the present review is concerned with the stability of children's mental health whilst growing up in care, three studies that employed a prospective timeframe of less than 12 months were excluded (Bogart, 1988; Damen & Pijnenburg, 2005; Portwood et al., 2018).

8. Population studies. To avoid conflating the effects of growing up in care with the effects of clinical or foster care interventions, only population (i.e. not treatment or intervention) studies were selected for review.

9. Mean baseline and follow-up scores reported. A 5-year prospective study was excluded because, rather than reporting mean scores from each of the 12-month follow-up assessments, it reported the mean of four annual follow-up mean scores (Kang, Woo, Chun, Nho, & Chung, 2017).

Statistical approaches

Aggregate change. Most of the studies selected for review reported group mean scores and standard deviations at two or more time points, from which mean change scores can be calculated. The present review reports mean change scores as standardized mean score differences (Cohen's *d*) – the difference in mean scores expressed as a proportion of the score standard deviation for the aggregate sample. Around half of these studies reported mean *raw* scores, and the remainder reported mean standardized *T* scores. Whereas raw scores are more precise than *T* scores, age-standardized *T* scores account for normative, developmental shifts in the population distribution of mental health difficulties, and thus provide a more appropriate metric for longitudinal studies that span from pre-adolescence through to adolescence.

Within-subject change. While prospective changes in group mean scores estimate aggregate change, they don't give a sense of how many children experience *meaningful* improvement or deterioration, or no meaningful change in their mental health. This information is fundamental to understanding the therapeutic and harmful effects of growing up in care. In practice, since children's mental health difficulties are experienced as continuously distributed phenomena, allocating change scores to change categories will always be imperfect. In doing so, it is important to distinguish between statistical significance, clinical significance, and a level of change that is perceptible to children and/or families, and/or that has developmental and social meaning.

Some studies have defined meaningful change as scores shifting from one severity range to another over time, such as from 'normal' to 'clinical', or from 'borderline' to 'normal'. One of the reviewed studies identified change trajectories from repeated measures obtained over an 8-year timeframe using a growth modelling procedure (Proctor et al., 2010). While this is useful for identifying changes in the rates of children who require clinical interventions, it is otherwise an inferior method for estimating meaningful change. This is because small, imperceptible changes in symptomatology (as little as a single raw score point) can push scores across a single clinical threshold (e.g., from borderline to normal).

Two of the reviewed studies reported rates of meaningful change using the Reliable Change Index (RCI). It describes the magnitude of change that is statistically reliable i.e. larger than what can be statistically attributable to internal measurement error (Jacobson & Truax, 1991). Jacobson and Truax (1991) originally proposed that the RCI be used as part of a 'twofold' procedure for establishing a level of change that is *clinically significant*, without elaborating on

what other criterion might be used to define clinically significant change. In practice, however, the RCI has since been employed in treatment evaluation as the *sole* criterion for defining clinically significant change. This approach is problematic, since there is no logical connection between the scale of score differences that is accounted for by measurement error, and that which represents clinically and developmentally meaningful change.

With this in mind, two of the reviewed studies attempted to define change that is *clinically meaningful*, and which is thus likely to be perceptible and meaningful to children and their caregivers. The first defined change in CBCL scores as meaningful if they were: 1. statistically significant (using the RCI); *and also* 2. clinically significant, as defined by a shift from the normal range to either the borderline or clinical range, or from the borderline or clinical range to the normal range (Vanderfaellie, Van Holen, Vanschoonlandt, Robberechts, & Stroobants, 2013). In clinical research, measuring change is based on the assumption that all study participants have clinical-level difficulties requiring intervention. However, population studies include participants who manifest normative mental health at baseline. Furthermore, the lower a child's symptom scores are at baseline, the less scope they have to experience improvement in their mental health – there is a 'floor' effect. For these children, a lack of meaningful change equates to sustained mental health. With this in mind, the second of the reviewed studies (Tarren-Sweeney, 2017) differentiated between those children whose scores were in the normal ranges on both occasions, and other participants, with the former constituting a *sustained mental health* group. For the latter group, meaningful change was then defined by the spread of scores traversing *both* borderline and clinical range cut-points. The reasoning for this was that a shift in scores from a normal range to a clinical range, and vice-versa, which traverses the borderline range, is clinically meaningful. This method yielded more conservative estimates of meaningful change than the RCI method (Tarren-Sweeney, 2017).

Review

Thirty-five longitudinal studies that measured prospective changes in the mental health of children in OOHC were located in the literature search. Of these, 21 studies were excluded from the review. These are listed in Table 1, together with reason(s) for their exclusion. Of the fourteen studies selected for the present review, ten measured children's mental health over a short timeframe (≤ 3 years) and four were long-term studies (≥ 5 years). The child participants in these fourteen studies ranged in age from 2 and 18 years. Two of the studies reported aggregated data for children residing in various types of care, but were included in the review because more than 90% were in family-based care.

[Insert Table 1 about here]

Question 1: What evidence is there that long-term, family-based care has a general, population-wide effect on children’s mental health?

If growing up in out-of-home care has a general, population-wide effect on children’s mental health (i.e. where the experience is generally therapeutic, or is generally harmful), it should manifest as a fairly constant mean rate of change across an entire population of children in care. Thus, assuming that the aggregate developmental impact that care systems exert on a population remains the same over time, the level of change measured over one year should be one fifth of that measured over 5 years. As time proceeds, however, longitudinal samples of children in care become progressively less representative, due to high sample attrition and various survivor biases, as well as their increasing age profile. Notwithstanding this critical limitation, it is useful to compare prospective mental health data in relation to the timeframes over which change was measured. Cohorts that are representative of care populations, particularly with respect to the distributions of ‘age at entry into care’, ‘time in care’ and ‘time in placement’, offer the possibility of estimating aggregate mental health changes over defined periods.

Table 2 lists estimates of *short-term* stability and change from five studies that recruited representative population samples, without respect to age at entry into care, time in care, and time in placement. Three of the studies (which include the two largest) measured very small or negligible ($d = -.06$ to $.09$) 12-month changes in: mean raw CBCL broadband scale scores among 8-13 year-old Dutch foster children ($N=53$) (Bastiaensen, 2001); mean caregiver-reported SDQ internalizing and externalizing difficulties among 4-17 year-old Dutch foster children ($N=180$) (Goemans, van Geel, & Vedder, 2018); and mean self-reported scores on scales measuring prosocial behaviour, emotional disorder and anxiety, conduct disorder and physical aggression, and relational aggression, among 10-17 year-old Canadian foster children ($N=201$) (Perkins, 2008).

The fourth study measured 12-month changes in socio-emotional difficulties and competence among 56 Norwegian 2-year-olds in foster care (Jacobsen, Moe, Ivarsson, Wentzel-Larsen, & Smith, 2013), using the Infant-Toddler Social and Emotional Assessment (ITSEA) (Carter, Briggs-Gowan, Jones, & Little, 2003). While baseline mean social-emotional difficulties T scores were unexpectedly low for children in care, so too were T scores for a Norwegian comparison group, suggesting the T score distributions (based on U.S. norms) are not valid for Norwegian toddlers. The foster children’s mean baseline social emotional difficulties T scores were only four to eight points higher than those for the comparison group. They manifested small to moderate 12-month increases in mean carer-reported externalizing ($d=0.17$) and internalizing ($d=0.40$) difficulties, contrasting with a small improvement in their mean competence scores ($d=0.18$), while mean dysregulation difficulties remained the same. However, the community sample’s mean internalizing difficulties increased by the same number of T score points ($d = 0.46$).

The fifth study reported rates of statistical change (RCI) over a 2-year period for 49 pre-adolescent foster children (Vanderfaellie et al., 2013). Eight (16%) showed statistical improvement in CBCL total problems scores, 18 (37%) showed deterioration, and 23 (47%) showed no change. Furthermore, the rate of children with CBCL total problems T scores ≥ 60 (borderline range cut-point) increased from 24% to 41%. While the follow-up rate (41%) is reasonably consistent with previous research estimates for this population (Oswald et al., 2010), the baseline rate (24%) falls well short of those estimates. Without understanding why the baseline rate for the 49 surviving participants was so low, it is difficult to interpret these findings.

[Insert Table 2 about here]

Table 3 lists estimates of *long-term* (≥ 5 years) stability and change from four studies that recruited representative population samples, with respect to age at entry into care, time in care, and time in placement. All of the long-term studies were afflicted by high sample attrition (33-75%), thereby limiting the interpretability of their findings, and highlighting the need for more definitive long-term cohort studies. The first study measured 5-year changes in depressive symptoms for 10-13 year-old (at baseline) Croatian children residing in long-term foster ($N=60$) (Bulat, 2010). The foster care sample had small reductions in mean self-reported depressive symptoms over five years, as measured by the YSR anxious-depressed sub-scale, and the Children's Depression Inventory (CDI), but a moderate increase in mean carer-reported CBCL anxious-depressed scores. While the self-report mean baseline scores are consistent with prior estimates of depressive symptoms for children in care, the mean carer-report score (2.73) is a little low (see for example, Simmel et al., 2014), potentially suggesting those carer-report scores may be unreliable.

The second study reported 7- to 9-year mental health changes for a small ($N=85$) sample of pre-adolescent (at baseline) Australian children in foster and kinship care (Tarren-Sweeney, 2017). The study showed no changes in mean CBCL age-standardized internalizing and externalizing T scores; small reductions in CBCL total problems ($d=0.20$) and CBCL Social-Attention-Thought problems (SAT) ($d=0.23$); and a small reduction ($d=0.26$) in attachment- and trauma-related difficulties, as measured by items common to the Assessment Checklist for Children (ACC) (baseline measure) and the Assessment Checklist for Adolescents (ACA) (follow-up measure) (the ACC-ACA score). The study also reported rates of meaningful change, as described in the method section. Around a third of the children manifested *sustained mental health* (35% based on the CBCL total score; 38% on the ACC-ACA score). Of the remaining 65% of children who had clinical or elevated total CBCL scores at baseline, roughly the same proportions ($\approx 40\%$) showed meaningful improvement, and meaningful deterioration, with the remaining 20% showing no meaningful change. The equivalent rates for those children who had clinical or elevated ACC-ACA scores at baseline were 40%, 30% and 30% respectively. While

this study had the longest prospective timeframe of the 14 studies, it also incurred the highest sample attrition (75%). However, analyses showed that those children retained at follow-up were broadly representative of the larger baseline sample ($N=347$).

The third study reported rates of meaningful change in carer-reported SDQ total difficulties scores over a 5-year period for a small ($N=60$) sample of pre-adolescent (at baseline) English foster children (Biehal, Ellison, Baker, & Sinclair, 2010), with similar rates of children manifesting ‘some or marked’ improvement (38%) and ‘some or marked’ deterioration (40%) and the remaining 22% showing no meaningful change.

The fourth study reported rates of statistical change (RCI) in CBCL total scores over an 8-year period for a small ($N=38$) sample of pre-adolescent (at baseline) Norwegian children in foster and kinship care (Vis, Handegård, Holtan, Fossum, & Thørnblad, 2016). Equal numbers of children ($N=10$, 26%) manifested statistically meaningful improvement and deterioration, while 18 (47%) showed no meaningful change.

[Insert Table 3 about here]

Conclusion on this question. While the evidence base is small and compromised by design limitations (notably high sample attrition), most studies that recruited representative population samples (with respect to age at entry into care, time in care, and time in placement) do not provide evidence that out-of-home care exerts a general, population-wide effect on the mental health of children in care. We conclude that there is no consistent evidence that growing up in care is generally reparative or generally harmful for children who enter care following exposure to severe social adversity.

Question 2: Does entry into long-term, family-based OOHC affect children’s mental health, as evidenced by prospective changes over the first years in care?

Table 4 lists estimates of *short-term* stability and change for the remaining five studies selected for review. The cohorts were not representative population samples, but instead were recruited in relation to their time in care or placements. Three studies recruited samples shortly after they entered care; one study recruited young children who had been in care for two years or less (Symanzik et al., 2019); and one study recruited a sample of ‘difficult to place’ children following placement with new foster families (Staines, 2012).

[Insert Table 4 about here]

Any effects that growing up in out-of-home care have on children’s mental health may not be uniform over time. Isolating non-linear, time-related effects is better achieved by following cohorts that are recruited at (or before) entry into care. Of the three studies that recruited cohorts at entry into care, two were separate cohorts in the U.S. National Survey of Child and Adolescent

Well-being (NSCAW), a nationally representative study conducted over five waves (baseline, 6-month, 18-month, 36-month, and 6- to 7-year follow-up) (Administration for Children Youth and Families, 2001). The NSCAW measured child and adolescent mental health from caregiver-reported CBCL scores, as well as self-reported scores on the Youth Self-Report (YSR, the self-report version of the CBCL) (age 11+), and post-traumatic stress sub-scale of the Trauma Symptom Checklist (TSCC) (age 8+).

The NSCAW Child Welfare (CW) cohort was 5,501 children aged 1-16 years at baseline, recruited to the study following child maltreatment notifications, including a sub-cohort who resided in out-of-home care at each stage of the study. Three published analyses of age-limited prospective data obtained for this sub-cohort are included in the present review. The first analysis compared 18- to 24-month mental health changes for 2-4 year-old children who: were placed in care ($N=152$); remained with their parents with support services ($N=274$); or remained with parents without services ($N=221$) (Stahmer et al., 2009). The in-care group had a sizeable, though non-significant reduction in mean caregiver-reported CBCL total problems T scores ($d = 0.44$), whereas children residing with their parents without support had increased scores ($d = 0.44$), and those receiving services had a smaller, non-significant increase. The second analysis reported 3-year changes in self-reported YSR internalizing and externalizing raw scores for 234 children between 11-14 years (Leonard & Gudiño, 2016). Mean baseline scores were within the range of previously reported estimates. Although an effect size could not be calculated (standard deviations not reported), participants reported a modest 3-year increase in mean externalizing scores, and a corresponding decrease in mean internalizing scores. The third analysis reported 18-month and 3-year changes in rates of CBCL disorders (internalizing and externalizing clinical ranges) for a wide-age sample (2-15 years at baseline, NSCAW) (Aarons et al., 2010). Baseline rates were within the range of previously reported estimates (Tarren-Sweeney & Hazell, 2006). The rate of externalizing disorders fell from 33.9% to 29.1% after 18 months, and then to 27.3% after 3 years, which is a sizeable reduction. The rate of internalizing disorders fell from 21.6% to 16.9% after 18 months, but rose to 20.8% after three years.

The NSCAW Long-Term Foster Care (LTFC) cohort consisted of 727 older children and adolescents who entered care approximately one year before baseline. At baseline, 91% of this sample resided in family-based care (58% in non-kin foster care, 32% in kinship care), and 9% were in group homes (i.e. small residential units) (Administration for Children Youth and Families, 2001). Three published analyses of age-limited prospective data obtained for this sub-cohort are included in the present review (Barboza, Dominguez, & Pinder, 2017; McWey, Cui, & Holtrop, 2014; McWey, Cui, & Pazdera, 2010). The first analysis identified small 18-month and 3-year reductions in carer-reported CBCL externalizing scores, and self-reported trauma symptom for 280 children between 8-15 years (Barboza et al., 2017). The second analysis identified small increases in self-reported YSR externalizing T scores at 18-month and 3-year follow-up, for 180

children between 11-16 years (McWey et al., 2014). However, baseline and follow-up mean externalizing scores are a little lower than expected for this population ($T=53-55$), suggesting the possibility that adolescents under-reported their difficulties. It is also not clear how the older adolescents were retained in the 3-year follow-up, by which time some would have been 19 years old. A parallel analysis of an older sub-set of this sample, namely 106 adolescents between 13-16 years (at baseline), identified 3-year changes in *carer-reported* CBCL externalizing and internalizing T scores (McWey et al., 2010). Baseline scores ($T=57, 62$) were within the range of previously reported estimates. There was a moderate 3-year reduction ($d=0.31$) in carer-reported externalizing scores, and a slight reduction in internalizing scores. Thus, adolescents and their carers reported 3-year mental health changes in opposing directions, with adolescents on average reporting modest deterioration, and their carers reporting slight to moderate improvement.

The third study measured 18-month mental health changes among 60 Dutch adolescents also contrasted carer-reported and self-reported scores (Strijker, van Oijen, & Knot-Dickscheit, 2011). In this study, young people reported moderate 18-month improvement in their mental health ($d=0.19-0.40$), while their carers reported moderate deterioration ($d=0.16-0.51$).

Conclusion on this question. The three studies present conflicting evidence on whether or not entry into out-of-home care has an *initial* effect on children's mental health, following their removal from maltreating families. The NSCAW is the best-designed longitudinal study of the mental health of children in care carried out to date. Analyses of carer-reported scores for various age ranges of the NSCAW samples identify small to moderate improvements in children's mean mental health scores during their first three years in care. However, the analysis by Aarons et al. (2010) suggests that while children may generally benefit emotionally from being removed from abusive care, for some children this effect is not sustained over the longer term. Conversely, in both NSCAW cohorts adolescent self-reported mean internalizing and externalizing difficulties increased slightly over the same timeframes.

Question 3: Is the reparative potential of long-term, family-based OOHC moderated by children's age at entry into care?

Several surveys have identified that older children and adolescents in care have greater mental health difficulties than younger children (Armsden, Pecora, Payne, & Szatkiewicz, 2000; Dubowitz, Zuravin, Starr, Feigelman, & Harrington, 1993; Heflinger, Simpkins, & Combs-Orme, 2000; Meltzer, Corbin, Gatward, Goodman, & Ford, 2003). This age effect is largely an artefact of later-placed children entering care with higher levels of pre-existing disturbance (Hukkanen, Sourander, Bergroth, & Piha, 1999b; Tarren-Sweeney, 2008). Nevertheless, while older age at entry into care is a marker for greater pre-care adversity, it might also moderate children's subsequent response and adjustment to care.

To what extent then do the prospective studies reviewed in this article shed light on the question of whether children's age at entry into care moderates the reparative potential of long-term care? Two of the studies measured changes in very young (i.e., 2-4 year-olds) children's socio-emotional development over short periods. The first study measured small to moderate mean *deterioration* in internalizing ($d=0.17$) and externalizing ($d=0.40$) difficulties (but improved competence, $d=0.18$) over 12 months among 2-year-olds in care (Jacobsen et al., 2013). The second measured a moderate mean *improvement* ($d=0.45$) over 18-24 months following entry into care (Stahmer et al., 2009). A third study measured modest one-year reductions in attachment disorder symptoms and interpersonal difficulties among a sample of two- to seven-year-old German children in foster care who had been in care for two years or less at baseline, and entered care following a history of maltreatment (Symanzik et al., 2019). However, neither of the carer-report measures used in this study were designed or validated for children under 4 or 5 years of age, and the study unexpectedly measured comparable one-year reductions in these symptoms and difficulties among a community sample of same-aged children. While this latter finding is difficult to interpret, it is possible that the measures that were designed for older children do not take account of normative relationship behaviours manifested by very young children.

Thus, while we know that younger age at entry into care predicts lower mental health difficulties, these studies do not clarify whether children who enter care at a young age are also more likely to experience improvement in their mental health. Similarly, longitudinal studies of the mental health trajectories of adolescents following their 'late arrival' into care yielded conflicting findings (Leonard & Gudiño, 2016; McWey et al., 2014; McWey et al., 2010; Strijker et al., 2011). The two studies that recruited representative adolescent foster care cohorts, measured slight 1-year reductions in mean self-reported difficulties (Perkins, 2008), and slight 5-year reduction in self-reported depressive symptoms (Bulat, 2010).

Conclusion on this question. None of the reviewed studies had sufficiently robust design, or adequate sample size and retention rates to definitively address this question. Similarly, there have not been enough prospective studies that recruited similar-age cohorts at entry into care to assess the consistency of any evidence.

Discussion

None of the research questions that we posed for this review are comprehensively answered by the available evidence. Perhaps our most important conclusion is that, as yet, no cohort study (or research programme) has had adequate design, scale or scope to provide a definitive understanding of the development and well-being of children who grow up in statutory care; or to contrast their developmental pathways with that of other high risk child populations growing up in different forms of care. Notwithstanding this uncertainty, the current research base

provides no evidence that out-of-home care exerts a general, population-wide effect on the mental health of children in care, consistent with Goemans et al.'s (2015) meta-analysis. In other words, they provide no evidence that growing up in care is generally reparative or generally harmful for children who enter care following exposure to severe social adversity. Instead, several longitudinal studies have demonstrated that sizeable proportions of children in care manifest meaningful improvement in their mental health over short- and long-term timeframes, and similarly sizeable proportions experience meaningful deterioration.

Various developmental theories (including attachment and social learning theories), as well as research into the neurodevelopmental effects of early maltreatment, would predict that the reparative and harm potentials of long-term care are moderated by such factors as children's age when entering care, their carers' commitment, the strength of their carers' relationships to them, and the stability of their placements. There are likely to be complex transactional mechanisms that shape children's developmental trajectories as they grow up in care. It is also important to understand that developmental change within care is moderated by children's earlier exposure to severe social adversity. The English and Romanian Adoption study found that the developmental effects of more than six months exposure to institutional deprivation in early childhood persists for many through childhood and adolescence – despite being subsequently raised by adoptive families (Sonuga-Barke et al., 2017). This supports the notion that recovery from some forms of psychopathology caused by early severe adversity tends to follow a long developmental trajectory even where a child's developmental conditions have markedly improved. There is even some evidence that early chronic maltreatment incurs a delayed 'sleeper effect' on later development, regardless of the quality of intervening care (Li & Godinet, 2014).

Therefore, rather than asking whether long-term care is generally beneficial or harmful for the development of previously maltreated children, future investigations should instead focus on the questions "*...what are the systemic and interpersonal characteristics of care that promote and sustain children's psychological development throughout childhood, and what characteristics are developmentally harmful?*" and "*...for which children is care beneficial, and for which children is it not?*" The answers to these questions are critical for improving policy and practice within children's services, and for designing more effective clinical interventions for this population. This knowledge will also help address the bigger question of whether our present out-of-home care systems can be remedied to the point that they adequately facilitate children's psychological development, or whether they should be abandoned. We know that large numbers and proportions of children placed into care effectively grow up without close and enduring familial relationships (Howard & Berzin, 2011; Reimer & Schäfer, 2015). Yet humans are a social species that evolved such that close and enduring familial relationships are essential for their psychosocial development. The absence of historical and ethnographic precedents for children growing up in impermanent caregiving systems (Boswell, 1988) infers this experience

lies outside the boundaries of human adaptation – in other words, that being raised without a semblance of a permanent family is both developmentally harmful, and contrary to human evolution.

As stated above, no research programmes have had adequate scale or scope to adequately address these questions. To do so will require large and ambitious cohort studies that overcome some major design obstacles, notably achieving adequate participant retention and reliable and valid measurement. Given children's dynamic care trajectories (including planned and unplanned placement changes, restoration to parental care, and shifts to permanent guardianship and adoption), adequate retention can only be feasibly attained by following children through various care arrangements, including restoration and permanent orders. This approach also offers scope for comparing the developmental trajectories of severely maltreated children who remain in their parents' care *versus* growing up in out-of-home care *versus* growing up in permanent guardianship / adoption *versus* subsequent restoration. However, broadening the scope in this way greatly increases the sample size required for these and other stratified analyses. Implementing any long-term cohort study of this type would also require considerable resourcing and expertise to sustain an acceptable participation rate, especially for children experiencing rapid placement changes. The method section of the present review highlights some critical limitations in measuring this population's mental health prospectively, including scope for systematic respondent biases, poor inter-rater reliability, and needing to employ different informants as children move through placements. Conducting a study that recruits information from different caregivers at different times in a child's life (such as parents, foster carers and adoptive parents) amplifies the risk of measurement error. With this in mind, we need to consider whether alternative, non-psychometric measures might yield additional, more accurate and reliable estimates of children's mental health in large-scale population cohort studies – including neurometric, biometric, and observational methods. Finally, given that many 'within care' experiences that have developmental significance are systemically-driven, and thus vary somewhat across child welfare jurisdictions, these questions need to be more definitively addressed through *cross-jurisdictional* and *cross-national* studies.

Table 1. Studies excluded from the review

Study	Location	Reason(s) for exclusion
Ahmed et al. (2005)	Iraq	Orphans placed in foster care
Barber & Delfabbro (2005)	Australia	Respondents were social workers
Berger et al. (2009)	U.S.	Mixed respondents: Baseline scores = birth parents, follow-up = foster carers
Bogart (1988)	U.S.	Small sample size (N<25); study interval < 12 months
Damen & Pijnenburg (2005); Damen & Veerman (2005)	Netherlands	Study interval < 12 months
Fanshel & Shin (1978); Frank (1980)	U.S.	Respondents were social workers
Fernandez (2008)	Australia	Not within-subjects analyses (unequal group N)
Gonzalez (1999)	U.S.	Small sample size (N<25)
Haight et al. (2010)	U.S.	Small sample size (N<25); Treatment study
Havnen et al. (2014)	Norway	Mixed respondents: Baseline scores = birth parents, follow-up = foster carers and birth parents.
Hiller & St. Clair (2018)	U.K.	Mean annual SDQ scores were reported for different combinations of the sample (large # missing values for each year). Therefore not within-subject comparisons.
Kang et al. (2017)	South Korea	Reported baseline mean scores, and the mean of the annual follow-up mean scores for years 2 through to 5
Lawrence et al. (2006)	U.S.	Small sample size (N<25)
Linares et al. (2007)	U.S.	Respondents were birth parents
Leathers et al. (2011)	U.S.	Small sample size (N<25)
McCauley and Trew (2000)	U.K.	Small sample size (N<25); study interval < 12 months
Minnis et al. (2006)	Scotland	Treatment study
LONGSCAN study: Newton et al. (2000); Proctor et al. (2010); Villodas et al. (2016)	U.S.	Not exclusively in care. Undisclosed number of children had returned to their parents
Portwood et al. (2018)	Canada	Study interval < 12 months
Rubin et al. (2007)	U.S.	Mixed respondents: Baseline scores = birth parents, follow-up = foster carers
Rushton, et al. (1995)	England	Small sample size (N<25)
White (1997)	U.S.	Small sample size (N<25)

Prospective mental health review

Table 2. Estimates of *short-term* stability and change: Representative population samples recruited without reference to age at entry into care, time in care, or time in placement

Study	Type of care	Baseline age range (years)	N	Attrition	Scale ^a	Follow-up interval (years)	Baseline Mean (SD)	Follow-up Mean (SD)	Effect size (Cohen's <i>d</i>)
Bastiaensen, 2001 (The Netherlands)	Foster	8-13	53	50%	CBCL <i>raw</i> total	2	35.5 (19.6)	33.3 (23.9)	-0.10
					CBCL <i>raw</i> ext.		11.9 (8.1)	12.0 (9.5)	+0.01
					CBCL <i>raw</i> int.		8.9 (6.2)	8.5 (6.8)	-0.06
Goemans et al., 2018 (The Netherlands)	Foster	4-17	180	58%	SDQ <i>raw</i> ext.	1	7.60 (4.64)	7.14 (4.40)	-0.10
					SDQ <i>raw</i> int.		5.03 (3.83)	5.02 (3.80)	0
Jacobsen et al., 2013 (Norway)	Foster	2	56	7%	ITSEA ext. <i>T</i>	1	52 (11.7)	54 (12.5)	+0.17
					ITSEA int. <i>T</i>		49 (10.2)	53 (9.9)	+0.40
					ITSEA dys. <i>T</i>		46 (10.4)	46 (13.7)	0
					ITSEA com. <i>T</i>		44 (11.7)	46 (10.4)	+0.18
Perkins, 2008 (Canada)	Foster	10-17	201	45%	AAR <i>raw</i> pro.	1	12.72 (4.05)	12.47 (4.10)	-0.06
					AAR <i>raw</i> emo.		4.84 (3.25)	4.70 (3.29)	-0.04
					AAR <i>raw</i> agg.		2.21 (2.35)	1.98 (2.44)	-0.09
					AAR <i>raw</i> relagg		2.20 (2.21)	2.02 (2.25)	-0.08
							Meaningful change rates ^b		
Vanderfaeillie et al., 2013 (Belgium)	Foster	6-12	49	36%	CBCL total <i>T</i>	2	<i>Improvement</i> 16.3%	<i>No change</i> 46.9%	<i>Deterioration</i> 36.7%
					CBCL ext. <i>T</i>		8.2%	61.2%	30.6%
					CBCL int. <i>T</i>		10.2%	73.5%	16.3%

^a ITSEA = Infant-Toddler Social and Emotional Assessment (carer-report); SDQ = Strengths and Difficulties Questionnaire (carer-report); AAR = Assessment and Action Record (AAR-C2) (self-report); CBCL = Child Behavior Checklist (carer-report); ext. = externalizing ; int. = internalizing; dys. = dysregulation; com. = competence; pro. = prosocial behaviour; emo. = emotional disorder and anxiety; agg. = conduct disorder and physical aggression; relagg. = indirect aggression (relational aggression); *T* = T-score; *raw* = raw score; behaviour

^b Meaningful change estimated by the Reliable Change Index

Table 3. Estimates of *long-term* stability and change: Representative population samples recruited without reference to age at entry into care, time in care, or time in placement

Study	Type of care	Baseline age range (years)	N	Attrition	Scale ^a	Follow-up interval (years)	Baseline Mean (SD)	Follow-up Mean (SD)	Effect size (Cohen's <i>d</i>)		
Bulat, 2010 (Croatia)	Foster	10-13	60	48%	CBCL anx dep.	5	2.73 (2.83)	3.71 (3.88)	+ 0.29		
					YSR anx-dep.		6.03 (4.06)	5.67 (4.96)	- 0.08		
					CDI total		8.32 (5.97)	7.53 (5.05)	- 0.14		
Tarren-Sweeney, 2017b (Australia)	Foster & kinship	4-11	85	75%	CBCL total <i>T</i>	7-9	59.4 (12.5)	56.9 (12.9)	- 0.20		
					CBCL ext. <i>T</i>		56.8 (12.1)	57.3 (12.3)	+ 0.04		
					CBCL int. <i>T</i>		52.7 (11.3)	52.7 (11.5)	0		
					CBCL SAT		17.2 (12.0)	14.4 (12.1)	- 0.23		
					ACC/ACA		17.0 (15.2)	14.4 (13.9)	- 0.26		
							Meaningful change rates				
							<i>Sustained^b mental health</i>	<i>Improvement</i>	<i>No change</i>	<i>Deterioration</i>	
							CBCL total	35.3%	27.1%	12.9%	24.7%
							ACC/ACA	37.7%	24.7%	18.8%	18.8%
							<i>Improvement^c</i>		<i>No change</i>	<i>Deterioration</i>	
							Marked	Some		Some	Marked
Biehal et al., 2010 (England)	Foster	4-11	60	33%	SDQ total	5	20%	18%	22%	25%	15%
							<i>Improvement^d</i>		<i>No change</i>	<i>Deterioration</i>	
Vis et al., 2016 (Norway)	Foster & kinship	4-9	38	52%	CBCL total	8	26%	48%	26%		

^a CBCL = Child Behavior Checklist (carer-report); YSR = Youth Self-Report (self-report); CDI = Children's Depression Inventory (self-report); SDQ = Strengths and Difficulties Questionnaire (carer-report); anx-dep. = anxious-depressed; ext. = externalizing ; int. = internalizing; SAT = nominal Social-Attention-Thought problems broadband scale; ACC/ACA = scale constructed from 64 items shared by the Assessment Checklist for Children (ACC) and Assessment Checklist for Adolescents (ACA); *T* = T-score.

^b 'Sustained mental health' defined as scores within normal range at baseline and follow-up. For all other scores: 'improvement' defined as CBCL total score reduction > 11, ACC-ACA shared-item score reduction > 4; 'no change' defined as CBCL total score change < 12, ACC-ACA shared-item score change < 5; 'deterioration' defined as CBCL total score increase > 11, ACC-ACA shared-item score increase > 4.

^c Change categories defined as: 'Marked improvement' = reduction of 5 or more points; 'Some improvement' = reduction of 2-4 points; 'No change' = < 2 point change; 'Some deterioration' = increase of 2-4 points; 'Marked deterioration' = increase of 5 or more points. No justification provided for the score ranges.

^d Meaningful change estimated by the Reliable Change Index (cut-off is 8 raw score points for the CBCL total problems scale)

Table 4. Estimates of *short-term* stability and change: Samples recruited with reference to *time in care* or *time in placement*

Study	Type of care	Analysis	Baseline age		Attrition	Scale ^a	Follow-up (years)	Mean scores							
			range (years)	N				Baseline Mean (SD)	Follow-up Mean (SD)	Effect Cohen's <i>d</i>					
1. Samples recruited following entry into care															
NSCAW, Cohort 1 (United States)	> 90% foster & kinship	Stahmer et al., 2009	2-4	152	Not stated	CBCL total <i>T</i>	1.5-2	56.8 (11.3)	51.6 (11.9)	-0.45					
			11-14	234			YSR <i>raw</i> ext.	3	14.7	16.6	SD not reported				
		Leonard & Gudiño, 2016				YSR <i>raw</i> int.		14.0	11.6						
								Clinical range rates							
								Baseline range rate	Follow-up range rate	Change					
		Aarons et al., 2010	2-15	500		CBCL ext. clinical	1.5	33.9%	29.1%	-4.8%					
						3		27.3%	-6.6%						
						1.5	21.6%	16.9%	-4.7%						
						3		20.8%	-0.8%						
								Baseline Mean (SD)	Follow-up Mean (SD)	Effect Cohen's <i>d</i>					
NSCAW, Cohort 2 (United States)	> 90% foster & kinship	Barboza et al., 2017	8-15	280	Unclear, but low	CBCL ext. <i>T</i>	1.5	60.1 (12.5)	59.4 (12.3)	-0.04					
							3		58.4 (13.2)	-0.13					
						TSCC-PTS <i>T</i>	1.5	48.0 (10.1)	47.4 (9.3)	-0.06					
			3		46.9 (9.5)	-0.11									
			McWey et al. 2014	11-16	180		YSR ext. <i>T</i>	1.5	53.3 (12.4)	54.7 (12.2)	+0.11				
								3		55.2 (10.4)	+0.17				
	McWey et al., 2010	13-16	106		CBCL ext. <i>T</i>	3	61.5 (11.6) ^b	57.8 (12.1)	-0.31						
					CBCL int. <i>T</i>		57.2 (12.9)	56.5 (10.9)	-0.06						
Strijker et al., 2011 (The Netherlands)	Foster & kinship		11-17	60	23%	CBCL <i>raw</i> total	1.5	30.4 (20.1)	39.9 (26.4)	+0.41					
						CBCL <i>raw</i> ext.		9.7 (7.5)	14.2 (10.3)	+0.51					
						CBCL <i>raw</i> int.		9.3 (7.1)	10.5 (8.0)	+0.16					
						YSR <i>raw</i> total		37.0 (18.4)	31.5 (16.4)	-0.32					
						YSR <i>raw</i> ext.		12.2 (6.4)	9.8 (5.7)	-0.40					
						YSR <i>raw</i> int.		10.7 (7.4)	9.4 (6.1)	-0.19					
2. Children in care for 2 years or less at baseline															
Symanzik et al, 2019 (Germany)	Foster		2-7	71	17%	ACC <i>raw</i> pseud.	1	1.97 (2.11)	1.71 (1.85)	-0.13	Comm. ^c				
						ACC <i>raw</i> inse.		3.50 (3.62)	2.89 (3.12)	-0.18	-0.21				
						ACC <i>raw</i> indis.		6.21 (3.46)	5.26 (3.12)	-0.29	-0.46				
									ACC <i>raw</i> total		11.68 (7.59)	9.87 (6.49)	-0.26	-0.35	
								72	16%	RPQ <i>raw</i> disin.		3.46 (3.62)	2.82 (3.46)	-0.18	-0.28
										RPQ <i>raw</i> inhib.		2.17 (2.53)	1.47 (2.70)	-0.27	-0.38
										RPQ <i>raw</i> total		5.64 (5.23)	4.29 (5.40)	-0.25	-0.39
3. Sample recruited following entry into new placement															
Staines, 2012 (England & Wales)	Foster		5-14	220	2%			Clinical range rates							
								Baseline range rate	Follow-up range rate	Change					
						SDQ normal	1	36%	37%	+1%					
						SDQ borderline		25%	20%	-5%					
					SDQ clinical		39%	43%	+4%						

^a CBCL = Child Behavior Checklist (carer-report); YSR = Youth Self-Report (self-report); ext. = externalizing ; int. = internalizing; T = T-score; raw = raw score; TSCC PTS = Post-traumatic stress sub-scale of the Trauma Symptom Checklist for Children (self-report); ACC = Assessment Checklist for Children – Short Form; pseud. = pseudomature interpersonal behaviour; inse. = insecure interpersonal behaviour; indis. = indiscriminate interpersonal behaviour; RPQ = Relationship Problems Questionnaire; disin. = disinhibited behaviour; inhib. = inhibited; SDQ = Strengths and Difficulties Questionnaire (carer-report); SDQ ‘normal’, ‘borderline’ and ‘clinical’ refer to ranges

^b Aggregate group mean scores and standard deviations estimated from means and standard deviations reported separately for each gender. The estimations are approximate.

^c Effect size (Cohen's *d*) for community comparison sample (n = 128-131)

References

- Aarons, G., James, S., Monn, A., Raghavan, R., Wells, R., & Leslie, L. (2010). Behavior problems and placement change in a national child welfare sample: A prospective study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(1), 70-80.
- Administration for Children Youth and Families. (2001). *National Survey of Child and Adolescent Well-being: One year in foster care report*. Washington, D.C.:
- Armsden, G., Pecora, P. J., Payne, V. H., & Szatkiewicz, J. P. (2000). Children placed in long-term foster care: An intake profile using the child behavior checklist/4-18 [yes]. *Journal of Emotional & Behavioral Disorders*, 8(1), 49-64.
- Barber, J., & Delfabbro, P. (2005). Children's adjustment to long-term foster care. *Children and Youth Services Review*, 27, 329-340.
- Barboza, G., Dominguez, S., & Pinder, J. (2017). Trajectories of post-traumatic stress and externalizing psychopathology among maltreated foster care youth: A parallel process latent growth curve model. *Child Abuse & Neglect*, 72, 370-382.
- Bastiaansen, P. (2001). *Belaste pleegouders en verscheurde pleegkinderen. Onderzoek naar de wijze waarop pleegouders de opvoedingssituatie belevenen pleegkinderen de relatie met het pleeggezin en het gezin van oorsprong beleven*. Zundert, The Netherlands: Vorsselmans.
- Berger, L., Bruch, S., Johnson, E., James, S., & Rubin, D. (2009). Estimating the impact of out-of-home placement on child well-being: Approaching the problem of selection bias. *Child Development*, 80(6), 1856-1876.
- Biehal, N., Ellison, S., Baker, C., & Sinclair, I. (2010). *Belonging and permanence: Outcomes in long-term foster care and adoption*. London: BAAF.
- Bogart, N. (1988). *A comparative study of behavioral adjustment between therapeutic and regular foster care in the treatment of child abuse and neglect*, Memphis State University, Memphis, TN).
- Boswell, J. (1988). *The kindness of strangers: The abandonment of children in western europe from late antiquity to the renaissance*. Chicago: University of Chicago Press.
- Bowlby, J. (1988). *A secure base: Clinical applications of attachment theory*. London: Routledge.
- Bulat, L. (2010). A longitudinal study of depressiveness in children in public care. *International Journal of Social Welfare*, 19(412-423).
- Burge, P. (2007). Prevalence of mental disorders and associated service variables among Ontario children who are permanent wards. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, 52(5), 305-314.
- Carter, A., Briggs-Gowan, M., Jones, S., & Little, T. (2003). The Infant-Toddler Social and Emotional Assessment (ITSEA): Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology*, 31(5), 495-514.
- Charlotte, C., Viding, E., Fearon, P., Glaser, D., & McCrory, E. (2017). Disentangling the mental health impact of childhood abuse and neglect. *Child Abuse & Neglect*, 63, 106-119.
- Damen, H., & Pijnenburg, H. (2005). *Onbaatzuchtig en professioneel: Innovatie in de Brabantse pleegzorg: Eindrapport monitorproject VersterkingPleegzorg Noord-Brabant 2003-2004 Praktikon*.
- Delfabbro, P., & Barber, J. (2003). Before it's too late: Enhancing the early detection and prevention of long-term placement disruption [yes]. *Children Australia*, 28(2), 14-18.
- Dozier, M., Grasso, D., Lindheim, O., & Lewis, E. (2007). The role of caregiver commitment in foster care: Insights from the This Is My Baby Interview. In D. Oppenheim & D. Goldsmith (Eds.), *Attachment theory in clinical work with children: Bridging the gap between research and practice* (pp. 90-108). New York: Guilford Press.
- Dozier, M., Stovall, K. C., Albus, K. E., & Bates, B. (2001, Sep-Oct). Attachment for infants in foster care: The role of caregiver state of mind [yes]. *Child Development*, 72(5), 1467-1477.
- Dubowitz, H., Zuravin, S., Starr, R. H., Feigelman, S., & Harrington, D. (1993, Dec). Behavior problems of children in kinship care. *Journal of Developmental and Behavioral Pediatrics*, 14(6), 386-393.

- English, D. J., & Graham, J. C. (2000). An examination of relationships between children's protective services social worker assessment of risk and independent LONGSCAN measures of risk constructs. *Children & Youth Services Review, 22*(11-12), 897-933.
- Fanshel, D., & Shinn, E. (1978). *Children in foster care: A longitudinal investigation*. New York: Columbia University Press.
- Fernandez, E. (2008). Unravelling emotional, behavioural and educational outcomes in a longitudinal study of children in foster care. *British Journal of Social Work, 38*(1283-1301).
- Frank, G. (1980). Treatment needs of children in foster care. *American Journal of Orthopsychiatry, 50*(2), 256-263.
- Goemans, A., van Geel, M., & Vedder, P. (2015). Over three decades of longitudinal research on the development of foster children: A meta-analysis. *Child Abuse and Neglect, 42*, 121-134.
- Goemans, A., van Geel, M., & Vedder, P. (2018). Foster children's behavioral development and foster parent stress: testing a transactional model. *Journal of Child and Family Studies, 27*(3), 990-1001.
- Gonzalez, E. (1999). *Cognitive and behavioral/emotional development of latino children in foster care*, California School of Professional Psychology, Alameda, California).
- Haight, W., Black, J., & Sheridan, K. (2010). A mental health intervention for rural, foster children from methamphetamine-involved families: Experimental assessment with qualitative elaboration. *Children & Youth Services Review, 32*(10), 1446-1457.
- Havnen, K., Breivik, K., & Jakobsen, R. (2014). Stability and change: A 7- to 8-year follow-up study of mental health problems in Norwegian children in long-term out-of-home care. *Child & Family Social Work, 19*, 292-303.
- Heflinger, C. A., Simpkins, C. G., & Combs-Orme, T. (2000, Jan). Using the CBCL to determine the clinical status of children in state custody [yes]. *Children & Youth Services Review, 22*(1), 55-73.
- Hiller, R., & St. Clair, M. (2018). The emotional and behavioural symptom trajectories of children in long-term out-of-home care in an English local authority. *Child Abuse & Neglect, 81*, 106-117.
- Howard, J., & Berzin, S. (2011). Never too old: Achieving permanency and sustaining connections for older youth in foster care. Retrieved from https://www.adoptioninstitute.org/wp-content/uploads/2013/12/2011_07_21_NeverTooOld.pdf
- Hukkanen, R., Sourander, A., Bergroth, L., & Piha, J. (1999a). Follow-up of children and adolescents in residential care in children's homes. *Nordic Journal of Psychiatry, 53*(3), 185-189.
- Hukkanen, R., Sourander, A., Bergroth, L., & Piha, J. (1999b). Psychosocial factors and adequacy of services for children in children's homes. *European Child and Adolescent Psychiatry, 8*, 268-275.
- Jacobsen, H., Moe, V., Ivarsson, T., Wentzel-Larsen, T., & Smith, L. (2013). Cognitive development and social-emotional functioning in young foster children: A follow-up study from 2 to 3 years of age. *Child Psychiatry and Human Development, 44*, 666-677.
- Jacobson, N., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.
- Kang, H., Woo, S., Chun, J., Nho, C., & Chung, I. (2017). Developmental outcomes among Korean adolescents in out-of-home care: A longitudinal study comparing kinship foster care and institutional care. *Children & Youth Services Review, 76*, 118-124.
- Lawrence, C., Carlson, E., & Egeland, B. (2006). The impact of foster care on development [yes]. *Development & Psychopathology, 18*(1), 57-76.
- Leathers, S., Spielfogel, J., McMeel, L., & Atkins, M. (2011). Use of a parent management training intervention with urban foster parents: A pilot study. *Children & Youth Services Review, 33*(7), 1270-1279.
- Leonard, S., & Gudiño, O. (2016). Academic and mental health outcomes of youth placed in out-of-home care: The role of school stability and engagement. *Child Youth Care Forum, 45*, 807-827.
- Li, F., & Godinet, M. (2014). The impact of repeated maltreatment on behavioral trajectories from early childhood to early adolescence. *Children and Youth Services Review, 36*, 22-29.
- Linares, L., Li, M., Shrout, P., Brody, G., & Pettit, G. (2007). Placement shift, sibling relationship quality, and child outcomes in foster care: A controlled study. *Journal of Family Psychology, 21*(4), 736-743.

- McCrae, J., & Barth, R. (2008). Using cumulative risk to screen for mental health problems in child welfare. *Research on Social Work Practice, 18*(2), 144-159.
- McCrory, E., & Viding, E. (2015). The theory of latent vulnerability: Reconceptualizing the link between childhood maltreatment and psychiatric disorder. *Development & Psychopathology, 27*, 493-505.
- McWey, L., Cui, M., & Holtrop, K. (2014). Caregiver–adolescent dyadic disagreement and problem behaviors of adolescents in the child welfare system. *Journal of Research on Adolescence, 25*(1), 44-50.
- McWey, L., Cui, M., & Pazdera, A. (2010). Changes in externalizing and internalizing problems of adolescents in foster care. *Journal of Marriage and Family, 72*(1128-1140).
- Meltzer, H., Corbin, T., Gatward, R., Goodman, R., & Ford, T. (2003). *The mental health of young people looked after by local authorities in England*. London: Office for National Statistics, The Stationery Office.
- Milan, S. E., & Pinderhughes, E. E. (2000, Win). Factors influencing maltreated children's early adjustment in foster care [yes]. *Development and Psychopathology, 12*(1), 63-81.
- Newton, R. R., Litrownik, A. J., & Landsverk, J. A. (2000, Oct). Children and youth in foster care: Disentangling the relationship between problem behaviors and number of placements [yes]. *Child Abuse and Neglect, 24*(10), 1363-1374.
- Oswald, S., Heil, K., & Goldbeck, L. (2010). History of maltreatment and mental health problems in foster children: A review of the literature. *Journal of Pediatric Psychology, 35*(5), 462-472.
- Perkins, J. (2008). *Foster parenting practices as predictors of foster child outcomes*, University of Ottawa, Ottawa).
- Portwood, S., Boyd, S., Nelson, E., Murdock, T., Hamilton, J., & Miller, A. (2018). A comparison of outcomes for children and youth in foster and residential group care across agencies. *Children & Youth Services Review, 85*, 19-25.
- Proctor, L., Skriner, L., Roesch, S., & Litrownik, A. (2010). Trajectories of behavioral adjustment following early placement in foster care: Predicting stability and change over 8 years. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(5), 464-473.
- Quiroga, M., & Hamilton-Giachritsis, C. (2016). Attachment styles in children living in alternative care: A systematic review of the literature. *Child Youth Care Forum, 45*, 625-653.
- Reimer, D., & Schäfer, D. (2015). The use of biographical narratives to explain favourable and unfavourable outcomes for children in foster care. *Adoption & Fostering, 39*(1), 5-20.
- Rubin, D., O'Reilly, A., Luan, X., & Localio, R. (2007). The impact of placement stability on behavioral well-being for children in foster care. *Pediatrics, 119*(2), 336-344.
- Rushton, A., Treseder, J., & Quinton, D. (1995, May). An eight-year prospective study of older boys placed in permanent substitute families: A research note [yes]. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 36*(4), 687-695.
- Schofield, G. (2002). The significance of a secure base: A psychosocial model of long-term foster care [yes]. *Child & Family Social Work, 7*, 259-272.
- Sonuga-Barke, E., Kennedy, M., Kumsta, R., Knights, N., Golm, D., Rutter, M., . . . Kreppner, J. (2017). Child-to-adult neurodevelopmental and mental health trajectories after early life deprivation: the young adult follow-up of the longitudinal English and Romanian Adoptees Study. *The Lancet*. [https://doi.org/10.1016/S0140-6736\(17\)30045-4](https://doi.org/10.1016/S0140-6736(17)30045-4)
- Stahmer, A., Hurlburta, M., Horwitz, S., Landsverk, J., Zhang, J., & Leslie, L. (2009). Associations between intensity of child welfare involvement and child development among young children in child welfare. *Child Abuse & Neglect, 33*, 598-611.
- Staines, J. (2012). Fostered children's behavioural and emotional difficulties: Findings from one independent foster care agency. *Clínica y Salud, 23*(3), 205-219.
- Strijker, J., van Oijen, S., & Knot-Dickscheit, J. (2011). Assessment of problem behaviour by foster parents and their foster children. *Child & Family Social Work, 16*, 93-100.
- Symanzik, T., Lohaus, A., Job, A., Chodura, S., Konrad, K., Heinrichs, N., & Reindl, V. (2019). Stability and change of attachment disorder symptoms and interpersonal problems in foster children. *Mental Health & Prevention, 13*, 35-42.

- Tarren-Sweeney, M. (2008). Retrospective and concurrent predictors of the mental health of children in care. *Children & Youth Services Review, 30*, 1-25.
- Tarren-Sweeney, M. (2017). Rates of meaningful change in the mental health of children in long-term out-of-home care: A seven- to nine-year prospective study. *Child Abuse and Neglect, 72*, 1-9.
- Tarren-Sweeney, M., & Hazell, P. (2006). The mental health of children in foster and kinship care in New South Wales, Australia. *Journal of Paediatrics & Child Health, 42*, 91-99.
- van den Dries, L., Juffer, F., van IJzendoorn, M., & Bakermans-Kranenburg, M. (2009). Fostering security? A meta-analysis of attachment in adopted children. *Children and Youth Services Review, 31*(3), 410-421.
- Vanderfaeillie, J., Van Holen, F., Vanschoonlandt, F., Robberechts, M., & Stroobants, T. (2013). Children placed in long-term family foster care: A longitudinal study into the development of problem behavior and associated factors. *Children and Youth Services Review, 35*(4), 587-593.
- Villodas, M., Litrownik, A., Newton, R., & Davis, I. (2016). Long-term placement trajectories of children who were maltreated and entered the child welfare system at an early age: Consequences for physical and behavioral well-being. *Journal of Pediatric Psychology, 41*(1), 46-54.
- Vis, S., Handegård, B., Holtan, A., Fossum, S., & Thørnblad, R. (2016). Social functioning and mental health among children who have been living in kinship and non-kinship foster care: Results from an 8-year follow-up with a Norwegian sample. *Child & Family Social Work, 21*, 557-567.
- White, D. (1997). *Using social problem-solving interventions with foster children: A means of preventing later psychopathology*, California School of Professional Psychology, Fresno, California).