

Effectiveness studies (randomised controlled trials)***Flynn 2012***

Study type	Randomised controlled trial (RCT)
Study location	Canada
Study setting	Tutoring delivered by foster carers to children in foster care
Study dates	2008 to 2009
Duration of follow-up	Post intervention testing (unclear duration of follow up)
Sources of funding	the Canada Education Savings Plan, Human Resources and Skills Development Canada, Government of Canada
Inclusion criteria	<p>Age Aged 6-13 years and in grade 2-7</p> <p>Care situation foster or kinship care home; living in a placement assessed as stable by child welfare worker and supervisor; possessing the legal status of a Crown Ward or Society Ward</p> <p>Other nominated by their child welfare worker as likely to benefit;</p> <p>Language English speaking</p>
Exclusion criteria	<p>Care situation Living in group home</p> <p>Education level Either very strong students or extremely weak students</p> <p>Behavioural</p>

	very behaviourally disturbed
Sample size	77
Split between study groups	42 foster children randomised to the TYCW group and 35 to the wait list group
Loss to follow-up	12 were lost to follow up in the TYCW group and 1 in the wait list group
% Female	Not reported
Mean age (SD)	10.7 ± 1.6 years
Outcome measures	<p>Educational outcome 1</p> <p>Wide Range Achievement Test—Fourth Edition (WRAT4). a standardized, norm-referenced test that assesses basic reading and math skills. It was developed for use with individuals aged 5–94 or in Grades K12. The WRAT4 comprises four subtests, Word Reading, Sentence Comprehension, Spelling, and Math Computation, and also yields a Reading Composite score that is obtained by combining the Word Reading and Sentence Comprehension standard scores.</p>
Study arms	<p>Foster parent-delivered Teach Your Children Well tutoring (N = 30)</p> <p>The foster children in the experimental group received tutoring and a Registered Education Saving Plan (to encourage saving for secondary education). The TYCW tutoring intervention was designed to provide 3 h per week of individual tutoring, for 30 weeks. The 3 h of weekly tutoring was to consist of 2 h of one-on-one direct instruction to the foster child in reading, 30 min of reading aloud by the foster child to the tutor or another adult in the home, and 30 min of self-paced instruction in math for the foster child, under the supervision of the foster parent. The math component was taught through step-by-step instruction in the form of a computer-based CD-ROM that the foster child used at his or her own pace. The reading component consisted of a four-level learn-to-read series of books, written by the designer of the TYCW program, Michael Maloney, and his team. For each reading level, there was a detailed instructor's manual and a student reader, and, for some levels, a student workbook as well. To determine the level of the TYCW program at which the foster child was to begin, his or her current reading level was determined by means of a standard assessment passage, administered by a research team member immediately after the child had been randomly assigned to the tutoring or wait-list control group. Also, to promote behavioral self-regulation and optimal learning, the TYCW program incorporated a behavior-</p>

	<p>management component, based on a reward system in which the child was to be awarded points for positive behavior in a particular tutoring session.</p> <table border="1"> <tr> <td>% Female</td> <td>50%</td> </tr> <tr> <td>Mean age (SD)</td> <td>Not reported</td> </tr> <tr> <td>Outcome measures</td> <td> <p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 100.32 (p=0.19); Spelling: 97.67 (p=0.74); Maths: 92.10 (p=0.009); Sentence comprehension: 103.22 (p=0.035) Reading composite: 101.23 (p=0.096). Adjusted for pre-intervention means.</p> </td> </tr> </table>	% Female	50%	Mean age (SD)	Not reported	Outcome measures	<p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 100.32 (p=0.19); Spelling: 97.67 (p=0.74); Maths: 92.10 (p=0.009); Sentence comprehension: 103.22 (p=0.035) Reading composite: 101.23 (p=0.096). Adjusted for pre-intervention means.</p>
% Female	50%						
Mean age (SD)	Not reported						
Outcome measures	<p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 100.32 (p=0.19); Spelling: 97.67 (p=0.74); Maths: 92.10 (p=0.009); Sentence comprehension: 103.22 (p=0.035) Reading composite: 101.23 (p=0.096). Adjusted for pre-intervention means.</p>						
	<p>Wait list (N = 34) The control children received the TYCW tutoring intervention during the school year (2009–2010) following that in which the experimental children had been tutored (2008–2009). During both years, each of the foster children in the experimental and control groups received a Registered Education Saving Plan (RESP) from their respective CAS for future post-secondary educational purposes. (RESPs are financial instruments created by the Government of Canada to encourage families and organizations such as CASs to save for children's post-secondary education.) Each child was assured of having \$1400 deposited in his or her RESP account. The foster parents in the two groups agreed to communicate weekly or more often to their tutees that the RESP was a symbol of their value as persons and a concrete financial investment in their futures.</p> <table border="1"> <tr> <td>% Female</td> <td>57.1%</td> </tr> <tr> <td>Mean age (SD)</td> <td>Not reported</td> </tr> <tr> <td>Outcome measures</td> <td> <p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 97.78; Spelling: 98.87; Maths: 86.30; Sentence comprehension: 98.69; Reading composite: 97.44. Adjusted for pre-intervention means.</p> </td> </tr> </table>	% Female	57.1%	Mean age (SD)	Not reported	Outcome measures	<p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 97.78; Spelling: 98.87; Maths: 86.30; Sentence comprehension: 98.69; Reading composite: 97.44. Adjusted for pre-intervention means.</p>
% Female	57.1%						
Mean age (SD)	Not reported						
Outcome measures	<p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 97.78; Spelling: 98.87; Maths: 86.30; Sentence comprehension: 98.69; Reading composite: 97.44. Adjusted for pre-intervention means.</p>						
Risk of bias	Domain 1: Bias arising from the randomisation process						

<p>Risk of bias judgement for the randomisation process</p> <p>Some concerns</p> <p>(Few baseline variables reported. Unclear if allocation concealment.)</p> <p>Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)</p> <p>High</p> <p>(Unclear if deviations from intended intervention. Per-protocol analysis and >30% drop out on the intervention arm).</p> <p>Domain 3. Bias due to missing outcome data</p> <p>High</p> <p>(Large loss to follow up and unclear how much missing data otherwise. Missing data imputed but unclear how much and if appropriate method used.)</p> <p>Domain 4. Bias in measurement of the outcome</p> <p>Some concerns</p> <p>(Outcome assessors were likely unblinded and outcome may be influenced by knowledge of intervention received (but not likely))</p> <p>Domain 5. Bias in selection of the reported result</p> <p>Some concerns</p> <p>(Unclear and insufficient detail provided about certain aspects of conducting trial e.g. approach to loss to follow up).</p> <p>Overall bias and Directness</p> <p>Risk of bias judgement</p> <p>High</p> <p>Overall Directness</p> <p>This question has not yet been answered.</p>

Geenen 2012

Study type	Randomised controlled trial (RCT)
Study location	USA
Study setting	Youth in Foster Care
Study dates	Not reported (published 2013)
Duration of follow-up	9 month follow up
Sources of funding	Funded by the Institute of Educational Sciences, U.S. Department of Education.
Inclusion criteria	<p>Age In the freshman, sophomore, or junior year of high school</p> <p>Care situation In the state foster care system</p> <p>Educational status receiving special education services within an urban school district</p>
Exclusion criteria	<p>Care situation scheduled to move out of state</p> <p>Language Non-English speaking</p>
Sample size	133
Split between study groups	63 in the TAKE CHARGE intervention group, 60 in the usual care group

Loss to follow-up	10 were lost to follow up in total, unclear how loss to follow up varied between intervention groups
% Female	46.3
Mean age (SD)	15.49 ± 2.21 years
Condition specific characteristics	<p>At risk or victims of exploitation Physical abuse: 38.2%; Sexual abuse: 33.3%; Neglect: 27.6%</p> <p>Disabilities, speech or communication needs, or special education needs Intellectual disability: 8.1%; Learning disability: 26.8%; Speech disability: 14.6%; Physical disability: 1.6%; Autism: 3.25%;</p> <p>Non-white ethnicity 50.4%</p> <p>Care characteristics Non kinship: 82.1%; Kinship: 13.0%; group home: 4.9%; length of time in foster care (mean): 84.6 months</p> <p>Number of placement moves mean 7.1</p>
Outcome measures	<p>Educational outcome 1 Youth knowledge and engagement in educational planning: measured using The student, parent, and teacher versions of the Educational Planning Assessment</p> <p>Educational outcome 2 Postsecondary preparation: On the outcome survey, youth completed a checklist indicating activities they had performed in planning for college. In all, 10 postsecondary items included "talked with guidance counselor or teacher about going to college" and "visited colleges". Item sums were calculated for each category.</p> <p>Educational outcome 3 Career development: Information regarding key activities youth had engaged in around career exploration and preparation for employment was also gathered on the outcome survey. 7 career items included "talked with family members about my career interests" and "job shadowed someone in my career area." Item sums were calculated for each category.</p> <p>Educational outcome 4 Student self-attribution of accomplishments: To assess selfattribution of educational success, conceptualized as an essential element of self-determination, youth were asked to list all their educational accomplishments for the past 6 months and a total count was gathered at each time point.</p> <p>Agency outcome 1 Self-determination: Self-determination was assessed with the parent, student, and teacher versions of the AIR as well as by asking youth to describe their goals and accomplishments as respective indices of youths' future directedness and positive self-attribution,</p> <p>Emotional and behaviour outcomes 1</p>

	<p>Measured with the Teacher Report Form (TRF) and the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001), and Youth Self-Report YSR (Achenbach, 1991). These parallel measures include scales for withdrawn-depressed, anxious-depressed, delinquent, and aggressive behavior, as well as attention problems. Analyses focused on the Withdrawn-Depressed, Anxious-Depressed, and Somatic Complaints subscales.</p> <p>Educational outcome 5 Student identification of education goals: At each time point, youth were asked to list all of their educational goals for the upcoming year and a total count was taken, gauged to reflect students' self-directedness.</p> <p>Educational outcome 6 Hours spent doing homework</p>						
Study arms	<p>TAKE CHARGE intervention (N = 60)</p> <p>Youth participated in two components of TAKE CHARGE: (a) Individualized coaching in applying self-determination skills to achieve their educational and related goals and to participate in educational planning meetings and (b) group mentoring, where the youth and near-peer foster care alumni who had completed high school and were working or in college gathered for information sharing and peer support. Mentors were recruited from college campuses, nominations from caseworkers, and study participants from earlier waves. To ensure fidelity, all coaches completed formal training and observation, and they attended weekly meetings where they discussed their work with youth and received ongoing support. Coaches also completed weekly log sheets where they documented the activities they engaged in and the time spent with each participant. The mean number of coaching sessions over an approximate 9-month period was 30.5 (SD = 7.8) with youth participating in an average of 32.97 (SD = 8.71) coaching hours over the duration of the intervention. Coaches and youth typically met weekly for 60 to 90 min; 13 was the minimum number of coaching hours and 55 was the maximum; youth availability accounted for much of the variation in coaching hours. Typically, one third of coaching time was didactic (M = 9.05, SD = 3.4) and two thirds experiential (M = 23.9, SD = 7.1). Overall fidelity for 79 coaching elements across all waves was 90.68%. Youth were invited to participate in three mentoring workshops, and they attended an average of 1.79 workshops. Workshop topics selected by youth included leading your education planning meeting, postsecondary education, careers, transportation, and relationships.</p> <table border="1" data-bbox="452 1185 2027 1401"> <tr> <td data-bbox="452 1185 689 1254">Study type</td> <td data-bbox="689 1185 2027 1254">Randomised controlled trial (RCT)</td> </tr> <tr> <td data-bbox="452 1254 689 1324">Study location</td> <td data-bbox="689 1254 2027 1324">USA</td> </tr> <tr> <td data-bbox="452 1324 689 1401">Study setting</td> <td data-bbox="689 1324 2027 1401">Youth in Foster Care</td> </tr> </table>	Study type	Randomised controlled trial (RCT)	Study location	USA	Study setting	Youth in Foster Care
Study type	Randomised controlled trial (RCT)						
Study location	USA						
Study setting	Youth in Foster Care						

Study dates	Not reported (published 2013)
Duration of follow-up	9 month follow up
Sources of funding	Funded by the Institute of Educational Sciences, U.S. Department of Education.
Inclusion criteria	Age In the freshman, sophomore, or junior year of high school
	Care situation In the state foster care system
	Educational status receiving special education services within an urban school district
Sample size	133
Split between study groups	63 in the TAKE CHARGE intervention group, 60 in the usual care group
Loss to follow-up	10 were lost to follow up in total, unclear how loss to follow up varied between intervention groups
% Female	40.0
Mean age (SD)	mean 15.79 years
Condition specific characteristics	At risk or victims of exploitation Physical abuse: 45.0%; Sexual abuse: 26.7%; Neglect: 26.7%
	Disabilities, speech or communication needs, or special education needs Intellectual disability: 8.3%; Learning disability: 26.7%; Speech disability: 23.3%; Physical disability: 45.0%; Autism: 1.7%
	Non-white ethnicity 53.3%

	<p>Care characteristics Non kinship: 85.0%; Kinship: 11.7%; group home: 4.9%; length of time in foster care (mean): 84.6 months</p> <p>Number of placement moves mean 7.9</p> <p>Educational outcome 1 Educational Planning Assessment score (following intervention/9-month follow up): Student-reported: 26.10 ± 5.71/26.61 ± 6.99; Parent reported: 22.13 ± 7.31/22.62 ± 8.05; Teacher reported: 20.40 ± 7.95/20.88 ± 7.84</p> <p>Educational outcome 2 Postsecondary preparation score: mean 2.53 ± 0.92/2.58 ± 0.94</p> <p>Educational outcome 3 Career development mean score (postintervention/9-month follow up): 2.64 ± 0.97/2.18 ± 0.78</p> <p>Educational outcome 4 Student self-attribution of accomplishments mean score (post-intervention/9-month follow up): 2.75 ± 1.44/2.31 ± 1.34</p> <p>Agency outcome 1 AIR self-determination score (post-intervention/9-month follow up): 66.43 ± 8.90/65.76 ± 8.56</p> <p>Emotional and behaviour outcomes 1 Youth Self Report Anxiety mean score (post-intervention/9-month follow up): 53.60 ± 5.11/54.09 ± 6.05; Child Behaviour Checklist anxiety: 55.33 ± 6.84/56.20 ± 6.94; Child Behaviour Checklist withdrawn score: 58.89 ± 7.04/58.23 ± 6.52; Child Behaviour Checklist somatic mean score: 57.84 ± 9.88/55.56 ± 6.52</p> <p>Educational outcome 5 Student identification of education goals score (postintervention/9-month follow up): 2.30 ± 1.23/1.90 ± 1.03</p> <p>Educational outcome 6 Hours spent doing homework mean (post intervention/9-month follow up): 1.32 ± 1.27/1.08 ± 1.13</p>
	<p>Outcome measures</p>
	<p>Usual Care (N = 60) Youth participating in the control group received typical educational services (business as usual), including general and special education classes, related services, interaction with special education case managers, individualized educational planning, and extracurricular activities.</p>

	<p>Outcome measures</p> <p>Educational outcome 1 Educational Planning Assessment score (following intervention/9-month follow up): Student-reported: 23.65 ± 7.85/23.93 ± 9.15; Parent reported: 19.32 ± 12.89/19.40 ± 8.14; Teacher reported: 17.89 ± 8.05/18.11 ± 8.90</p> <p>Educational outcome 2 Postsecondary preparation score (postintervention/9-month follow up): mean 1.52 ± 0.40/2.56 ± 0.89</p> <p>Educational outcome 3 Career development mean score (postintervention/9-month follow up): 2.04 ± 0.71/2.01 ± 0.69</p> <p>Educational outcome 4 Student self-attribution of accomplishments mean score (post-intervention/9-month follow up): 1.95 ± 1.20/2.07 ± 1.23</p> <p>Agency outcome 1 Parent reported AIR self-determination score (post-intervention/9-month follow up): 63.52 ± 8.94/62.96 ± 8.81</p> <p>Emotional and behaviour outcomes 1 Youth Self Report Anxiety mean score (post-intervention/9-month follow up): 56.19 ± 6.61/54.61 ± 5.79; Child Behaviour Checklist anxiety: 60.43 ± 8.60/59.00 ± 8.58; Child Behaviour Checklist withdrawn score: 62.36 ± 9.60/61.19 ± 9.08; Child Behaviour Checklist somatic mean score: 60.70 ± 9.39/60.00 ± 9.53</p> <p>Educational outcome 5 Student identification of education goals score (postintervention/9-month follow up): 2.05 ± 1.14/1.92 ± 1.05</p> <p>Educational outcome 6 Hours spent doing homework mean (post intervention/9-month follow up): 0.81 ± 1.11/0.94 ± 0.96</p>
<p>Risk of bias</p>	<p>Domain 1: Bias arising from the randomisation process</p> <p>High</p> <p>(Some considerable differences between comparison groups for length of time in foster care, speech and language disability, autism, and emotional/behavioural needs)</p> <p>Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)</p> <p>Some concerns</p> <p>(unclear if any deviations from intended interventions; unclear if intention to treat analysis used (but most likely))</p> <p>Domain 3. Bias due to missing outcome data</p>

	<p>High</p> <p>(Just over 10% with missing data post randomisation; unclear whether any further missing outcome data; unclear reasons for drop out; unclear how drop out varied between groups; It is possible that missingness of data is related to outcomes.)</p> <p>Domain 4. Bias in measurement of the outcome</p> <p>Some concerns</p> <p>(It is unclear how assessments were performed (by whom). Unclear if facilitators were aware of intervention status of participants. Measurements used are often crude indicators of the phenomenon of interest.)</p> <p>Domain 5. Bias in selection of the reported result</p> <p>High</p> <p>(unclear that analysis was conducted according to a pre-specified protocol. Data not provided for certain non-significant results. Evidence of multiple analyses used for different outcomes)</p> <p>Overall bias and Directness</p> <p>Risk of bias judgement</p> <p>High</p> <p>Overall Directness</p> <p>This question has not yet been answered.</p>
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Green 2014

Study type	Randomised controlled trial (RCT)
Study location	UK England
Study setting	Looked after young people (on a placement at risk of breakdown)
Study dates	June 2005 to December 2008

Duration of follow-up	12 months
Sources of funding	The project was funded by a grant from the UK Department for Children, Schools and Families to the Institute of Psychiatry (reference: ACLBMC). It was sponsored by the University of Manchester.
Inclusion criteria	<p>Age aged 10-17 years</p> <p>Care situation in a placement that was unstable, at risk of breakdown or not meeting their assessed needs, or at risk of custody or secure care</p> <p>Emotional or behavioral disorders showing complex or severe emotional difficulties and/or challenging behaviour</p>
Exclusion criteria	<p>Special educational needs severe intellectual difficulties (referred to as learning disabilities by UK health services, this was indexed by specialist school placement)</p> <p>Medical health problem psychotic illness from medical records.</p>
Sample size	34
Split between study groups	20 randomised to MTFC-A, 14 randomised to usual care
Loss to follow-up	3 lost to follow up in the MTFC-A group, 2 in the usual care group
% Female	Not reported for total population
Mean age (SD)	Not reported for total population
Outcome measures	<p>Global health outcome 1</p> <p>Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA): Sources included structured interviews with the young person and carers, the standard carer-rated Child Behaviour Checklist (CBCL) and self-rated Youth Self Report (YSR), 10 along with collated reports and records directly accessed from education, health and social services. This information was integrated, transcribed, fully anonymised and then located within each relevant HOTN domain before being rated. A second researcher, masked to all other case data including the first rating, independently rated this anonymised information within each domain.</p>

	<p>Global health outcome 2 Children's Global Assessment Scale (CGAS). Sources included structured interviews with the young person and carers, the standard carer-rated Child Behaviour Checklist (CBCL) and self-rated Youth Self Report (YSR),10 along with collated reports and records directly accessed from education, health and social services. This information was integrated, transcribed, fully anonymised and then located within each relevant CGAS domain before being rated. A second researcher, masked to all other case data including the first rating, independently rated this anonymised information within each domain.</p> <p>Educational outcome 1 Scholastic/language skills. Education outcomes were assessed using masked ratings on the two education-related HoNOSCA domains (scholastic/language skills and education attendance).</p> <p>Educational outcome 2 School attendance. Education outcomes were assessed using masked ratings on the two education-related HoNOSCA domains (scholastic/language skills and education attendance).</p> <p>Criminal outcome 1 Offending at follow up. Data on specific incidents of offending (reprimand, caution or charged with offence) during the previous 6 months were gathered from the social worker at baseline and from carer and social worker at end-point covering the previous 3 months.</p>
Study arms	<p>Multidimensional treatment foster care for adolescents (MTFC-A) (N = 20) In MTFC-A, specialist foster parents receive training and ongoing support and supervision in an intensive social learning approach pioneered at the Oregon Social Learning Center. Attention is paid to the mental health of foster children through the provision of psychiatry and psychology input, including individual and family therapy, social skills training and support with education. The aim is for a short-term intensive placement, of around 9 months, followed by a short period of aftercare. Key elements include: the provision of a consistent reinforcing environment in which young people are mentored and encouraged; a clear structure, with clearly specified boundaries to behaviour and specified consequences that can be delivered in a teaching-oriented manner; close supervision of young people's activities and whereabouts at all times; diversion from associations with antisocial peers and help to develop positive social skills that will help young people form relationships with more positive peers. Behaviour is closely monitored and positive behaviours are reinforced in a concrete manner using a system of points and levels; moving during the course of the programme from early restrictions through a series of 'levels,' each of which brings increased privileges and enhanced incentives. Specialist foster carers are paid a full-time salary, provided with continuously available intensive support, have daily telephone interviews with MTFC-A staff for support and to complete a Parent Daily Report (PDR), a checklist enabling the team to monitor intervention adherence, and identify problems, progress and carer stress. Foster carers have weekly face-to-face group meetings with the intervention team. Participating intervention teams received initial training from the UK national implementation group and the programme developers in the USA to prespecified levels of fidelity. Following this, ongoing fidelity to the model throughout the programme was monitored through weekly supervision telephone calls with the programme developers in</p>

<p>the USA, including evaluation of individual PDR data. In each local team there were two additions to the US model: (a) an education worker; and (b) a part-time programme manager to liaise with the Social Services department.</p>	
% Female	Not reported for RCT sample
Mean age (SD)	Not reported for RCT sample
Outcome measures	<p>Global health outcome 1 Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA) at 12 months: mean 14.04 ± 5.57. Adjusted mean difference between MTFC-A and usual care at follow up: -1.04 (-6.21 to 4.13). Adjusted for baseline score.</p>
	<p>Global health outcome 2 Children's Global Assessment Scale (CGAS) at 12 month follow up: mean 56.00 ± 10.06. Adjusted mean difference between MTFC-A and usual care at 12 months: 1.30 (-7.14 to 9.74). Adjusted for baseline score.</p>
	<p>Educational outcome 1 Scholastic/language skills. Odds of higher follow up score in the MTFC compared to usual care intervention group: OR 0.6 (95%CI 0.15 to 2.4)</p>
	<p>Educational outcome 2 School attendance. Odds of higher school attendance score in the MTFC group: 2.5 (95%CI 0.48 to 13.1)</p>
	<p>Criminal outcome 1 Number offending at follow up: 7. adjusted odds of offending in MTFC compared to usual care: aOR 1.24 (95%CI 0.22 to 7.38). Odds ratio adjusted for baseline offending age, gender, baseline offending and antisocial behaviour with inverse probability weighting by propensity score.</p>
<p>Usual care (N = 14) Usual care consisted of care placements routinely in use in local authorities at the time. These included existing (non-MTFC-A) family foster care, residential care, residential schools and other placements. Details of the use of these placements and of other mental health services were gathered at carer interview.</p>	
% Female	Not reported for RCT population
Mean age (SD)	Not reported for RCT population

	<table border="1"> <tr> <td data-bbox="448 276 689 547">Outcome measures</td> <td data-bbox="689 276 2042 547"> <p>Global health outcome 1 Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA) at 12 months follow up: mean score 14.93 ± 7.99</p> <p>Global health outcome 2 Children's Global Assessment Scale (CGAS) at 12 months follow up: mean score 55.25 ± 12.56</p> <p>Criminal outcome 1 Participants offending at follow up: 4</p> </td> </tr> </table>	Outcome measures	<p>Global health outcome 1 Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA) at 12 months follow up: mean score 14.93 ± 7.99</p> <p>Global health outcome 2 Children's Global Assessment Scale (CGAS) at 12 months follow up: mean score 55.25 ± 12.56</p> <p>Criminal outcome 1 Participants offending at follow up: 4</p>
Outcome measures	<p>Global health outcome 1 Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA) at 12 months follow up: mean score 14.93 ± 7.99</p> <p>Global health outcome 2 Children's Global Assessment Scale (CGAS) at 12 months follow up: mean score 55.25 ± 12.56</p> <p>Criminal outcome 1 Participants offending at follow up: 4</p>		
Risk of bias	<p>Domain 1: Bias arising from the randomisation process</p> <p>Low</p> <p>Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)</p> <p>Some concerns</p> <p>(Unclear if/why participants did not receive allocated intervention; Significant deviations apparent since 8/20 in the treatment group did not receive their interventions.)</p> <p>Domain 3. Bias due to missing outcome data</p> <p>High</p> <p>(In the intervention group 15-20% had missing data; it was also unclear how much other data was missing since some outcomes were imputed; Unclear if appropriate imputation methods used; reasons for missing data not given; Missingness of data may well be related to the result of the outcomes reported.)</p> <p>Domain 4. Bias in measurement of the outcome</p> <p>Low</p> <p>(However, outcomes were triangulated from multiple sources. Assessors were masked to treatment group.)</p> <p>Domain 5. Bias in selection of the reported result</p> <p>High</p> <p>Overall bias and Directness</p>		

	<p>Risk of bias judgement</p> <p>High</p> <p>Overall Directness</p> <p>This question has not yet been answered.</p>
--	-----------------------------------------------------------------------------------------------------------------------------------

Harper 2012

Study type	Randomised controlled trial (RCT)
Study location	Canada
Study setting	Small group tutoring programme for children in foster care
Study dates	September 2010 to April 2011
Duration of follow-up	Following the intervention (no specific length of follow up defined)
Sources of funding	funding provided through the Ministry of Children and Youth Services, Ministry of Education, and the Ministry of Training, Colleges and Universities.
Inclusion criteria	<p>Age Between grade 2 and 8</p> <p>Care situation Foster or kinship care</p> <p>Educational status behind in their academic achievement but not intellectually challenged (i.e. IQ>70)</p> <p>Other Able to remain in the study for the full 25 weeks of the intervention</p>
Exclusion criteria	<p>Special educational needs</p> <p>IQ <70</p>

Sample size	68
Split between study groups	33 randomised to TYCW and 35 to wait list
Loss to follow-up	3 lost to follow up in the TYCW group
% Female	42.6%
Mean age (SD)	Not reported
Condition specific characteristics	Non-white ethnicity 73.5% aboriginal ethnicity
Outcome measures	<p>Educational outcome 1</p> <p>Wide Range Achievement Test Fourth Edition (WRAT-4). The WRAT-4 was used to measure academic achievement across four dimensions including reading, spelling, sentence comprehension and mathematics. The WRAT-4 is standardized and norm-referenced, with all scores converted to standard scores with a mean of 100, and a standard deviation of 15. Word Reading measures letter and word decoding by word recognition and identification. Spelling measures the ability to encode sounds into written form by use of a dictated spelling format containing both letters and words. Sentence Comprehension measures the ability to gain meaning from words and to understand and comprehend ideas and information within the sentences. Math Computation measures the ability to perform and execute mathematical computations by counting, identifying numbers, solving simple oral problems and calculating written math problems.</p>
Study arms	<p>Volunteer tutor-delivered Teach Your Children Well (TYCW) (N = 30)</p> <p>A tutoring intervention program, children were assessed on a measure of word fluency, used in the TYCW program, and placed into small tutoring groups of three or four children according to skill level. The group-based tutoring groups ran over a 25-week time frame, for 2 h each week, with either one or two tutor volunteers running each group. Each session followed the basic structure of Michael Maloney's TYCW curriculum, which uses direct instruction and behaviour management to improve the educational attainment of children. Volunteer university students were recruited to run the weekly tutoring program. Prior to working with the children, the tutors completed two full days of training with the tutoring developer, Michael Maloney. Tutors were required to collect performance data at each tutoring session. This data comprised the fidelity checks and consisted of sound fluency (e.g. number of sounds read from a list of sounds per 30 s), word fluency (e.g. number of words read from a list per 30 s), and story fluency (e.g. number of words read from a story in 1 min). This data was compiled into a weekly spreadsheet that was sent to Mr. Maloney. Throughout the course of the</p>

	<p>study, tutors had their performance monitored by Mr. Maloney who served as an ongoing consultant. All volunteers received an honorarium at the middle and end of the tutoring program.</p> <table border="1"> <tr> <td>% Female</td> <td>Not reported</td> </tr> <tr> <td>Mean age (SD)</td> <td>Not reported</td> </tr> <tr> <td>Outcome measures</td> <td> <p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 93.81 (p=0.002); Spelling: 99.68 (p=0.004); Maths: 82.89 (p=ns); Sentence comprehension: 92.61 (p=ns). Adjusted for pre-intervention means.</p> </td> </tr> </table> <p>Wait list (N = 35) No further description</p> <table border="1"> <tr> <td>% Female</td> <td>Not reported</td> </tr> <tr> <td>Mean age (SD)</td> <td>Not reported</td> </tr> <tr> <td>Outcome measures</td> <td> <p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores: Word reading: 89.36; Spelling: 91.79; Maths: 79.69; Sentence comprehension: 91.75 (adjusted for pre-intervention means).</p> </td> </tr> </table>	% Female	Not reported	Mean age (SD)	Not reported	Outcome measures	<p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 93.81 (p=0.002); Spelling: 99.68 (p=0.004); Maths: 82.89 (p=ns); Sentence comprehension: 92.61 (p=ns). Adjusted for pre-intervention means.</p>	% Female	Not reported	Mean age (SD)	Not reported	Outcome measures	<p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores: Word reading: 89.36; Spelling: 91.79; Maths: 79.69; Sentence comprehension: 91.75 (adjusted for pre-intervention means).</p>
% Female	Not reported												
Mean age (SD)	Not reported												
Outcome measures	<p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 93.81 (p=0.002); Spelling: 99.68 (p=0.004); Maths: 82.89 (p=ns); Sentence comprehension: 92.61 (p=ns). Adjusted for pre-intervention means.</p>												
% Female	Not reported												
Mean age (SD)	Not reported												
Outcome measures	<p>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores: Word reading: 89.36; Spelling: 91.79; Maths: 79.69; Sentence comprehension: 91.75 (adjusted for pre-intervention means).</p>												
Risk of bias	<p>Domain 1: Bias arising from the randomisation process</p> <p>Low</p> <p>(Centralised randomisation prior to direct contact with participants)</p> <p>Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)</p> <p>Some concerns</p> <p>(Unclear if deviations from intended intervention; unclear why loss to follow up; Per-protocol analysis; <10% lost to follow up.)</p>												

	<p>Domain 3. Bias due to missing outcome data</p> <p>Low</p> <p>Domain 4. Bias in measurement of the outcome</p> <p>Some concerns</p> <p>(Unclear if outcome assessors were aware of a participants intervention status. It is possible that such knowledge could have impacted results.)</p> <p>Domain 5. Bias in selection of the reported result</p> <p>High</p> <p>(Unclear that analysis was conducted with a pre-specified plan e.g. for multivariable analysis; some evidence that multiple analyses were performed but only one reported. Raw data not reported.)</p> <p>Overall bias and Directness</p> <p>High</p> <p>Overall Directness</p> <p>This question has not yet been answered.</p>
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Harper 2016

Study type	Randomised controlled trial (RCT)
Study location	Canada
Study setting	Small group tutoring programme for children in foster care
Study dates	2010 and 2011

Duration of follow-up	Following the intervention (no specific length of follow up defined)
Sources of funding	funding provided through the Ministry of Children and Youth Services, Ministry of Education, and the Ministry of Training, Colleges and Universities.
Inclusion criteria	<p>Age Between grade 1 and 8</p> <p>Care situation Foster or kinship care</p> <p>Educational status behind in their academic achievement but not intellectually challenged (i.e. IQ>70)</p> <p>Other Able to remain in the study for the full 25 weeks of the intervention</p>
Exclusion criteria	<p>Special educational needs IQ <70</p>
Sample size	101
Split between study groups	49 randomised to TYCW intervention, 51 randomised to wait-list control
Loss to follow-up	9 lost to follow up (4 in the TYCW group, 5 in the wait list group).
% Female	42.6%
Mean age (SD)	Not reported
Condition specific characteristics	<p>Non-white ethnicity 78.2% aboriginal</p>
Outcome measures	Educational outcome 1

	<p>Wide Range Achievement Test Fourth Edition (WRAT-4). The WRAT-4 was used to measure academic achievement across four dimensions including reading, spelling, sentence comprehension and mathematics. The WRAT-4 is standardized and norm-referenced, with all scores converted to standard scores with a mean of 100, and a standard deviation of 15. Word Reading measures letter and word decoding by word recognition and identification. Spelling measures the ability to encode sounds into written form by use of a dictated spelling format containing both letters and words. Sentence Comprehension measures the ability to gain meaning from words and to understand and comprehend ideas and information within the sentences. Math Computation measures the ability to perform and execute mathematical computations by counting, identifying numbers, solving simple oral problems and calculating written math problems.</p>								
<p>Study arms</p>	<p>Volunteer tutor-delivered Teach Your Children Well (TYCW) (N = 45) A tutoring intervention program, children were assessed on a measure of word fluency, used in the TYCW program, and placed into small tutoring groups of three or four children according to skill level. The group-based tutoring groups ran over a 25-week time frame, for 2 h each week, with either one or two tutor volunteers running each group. Each session followed the basic structure of Michael Maloney's TYCW curriculum, which uses direct instruction and behaviour management to improve the educational attainment of children. Volunteer university students were recruited to run the weekly tutoring program. Prior to working with the children, the tutors completed two full days of training with the tutoring developer, Michael Maloney. Tutors were required to collect performance data at each tutoring session. This data comprised the fidelity checks and consisted of sound fluency (e.g. number of sounds read from a list of sounds per 30 s), word fluency (e.g. number of words read from a list per 30 s), and story fluency (e.g. number of words read from a story in 1 min). This data was compiled into a weekly spreadsheet that was sent to Mr. Maloney. Throughout the course of the study, tutors had their performance monitored by Mr. Maloney who served as an ongoing consultant. All volunteers received an honorarium at the middle and end of the tutoring program.</p> <table border="1" data-bbox="452 962 2033 1230"> <tr> <td data-bbox="452 962 689 1038">% Female</td> <td data-bbox="689 962 2033 1038">Not reported</td> </tr> <tr> <td data-bbox="452 1038 689 1115">Mean age (SD)</td> <td data-bbox="689 1038 2033 1115">Not reported</td> </tr> <tr> <td data-bbox="452 1115 689 1230">Outcome measures</td> <td data-bbox="689 1115 2033 1230"> Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 93.62 (p<0.001); Spelling: 94.80 (p=0.02); Maths: 84.27 (p=0.044); Sentence comprehension: 92.78 (p=ns). Adjusted for pre-intervention means. </td> </tr> </table> <p>Wait list (N = 51) No further description</p> <table border="1" data-bbox="452 1369 2033 1445"> <tr> <td data-bbox="452 1369 689 1445">% Female</td> <td data-bbox="689 1369 2033 1445">Not reported</td> </tr> </table>	% Female	Not reported	Mean age (SD)	Not reported	Outcome measures	Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 93.62 (p<0.001); Spelling: 94.80 (p=0.02); Maths: 84.27 (p=0.044); Sentence comprehension: 92.78 (p=ns). Adjusted for pre-intervention means.	% Female	Not reported
% Female	Not reported								
Mean age (SD)	Not reported								
Outcome measures	Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores post-intervention. Word reading: 93.62 (p<0.001); Spelling: 94.80 (p=0.02); Maths: 84.27 (p=0.044); Sentence comprehension: 92.78 (p=ns). Adjusted for pre-intervention means.								
% Female	Not reported								

	<table border="1"> <tr> <td>Mean age (SD)</td> <td>Not reported</td> </tr> <tr> <td>Outcome measures</td> <td>Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores: Word reading: 88.98; Spelling: 91.61; Maths: 80.43; Sentence comprehension: 91.08 (adjusted for pre-intervention means).</td> </tr> </table>	Mean age (SD)	Not reported	Outcome measures	Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores: Word reading: 88.98; Spelling: 91.61; Maths: 80.43; Sentence comprehension: 91.08 (adjusted for pre-intervention means).
Mean age (SD)	Not reported				
Outcome measures	Educational outcome 1 Wide Range Achievement Test Fourth Edition (WRAT-4) adjusted mean scores: Word reading: 88.98; Spelling: 91.61; Maths: 80.43; Sentence comprehension: 91.08 (adjusted for pre-intervention means).				
Risk of bias	<p>Domain 1: Bias arising from the randomisation process</p> <p>Low</p> <p>(Centralised randomisation prior to direct contact with participants)</p> <p>Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)</p> <p>Some concerns</p> <p>Domain 3. Bias due to missing outcome data</p> <p>Low</p> <p>Domain 4. Bias in measurement of the outcome</p> <p>Some concerns</p> <p>(It is possible that the test could have been influenced by prior knowledge of intervention group. Unclear that assessment was masked.)</p> <p>Domain 5. Bias in selection of the reported result</p> <p>High</p> <p>(Unclear that analysis was performed according to a pre-specified plan; unclear when post-test took place; Some evidence of multiple analysis techniques used but only one reported. Unclear how covariates were defined. No raw data presented.)</p> <p>Overall bias and Directness</p> <p>Risk of bias judgement</p> <p>High</p>				

	Overall Directness
	This question has not yet been answered.

Hickey 2020**Study details**

Study type	Randomised controlled trial (RCT)
Study location	Canada
Study setting	Two local Children's Aid Societies (CASs) in Ontario: school-aged foster children in care
Study dates	Not reported
Duration of follow-up	Post-intervention
Sources of funding	University of Ottawa,
Inclusion criteria	Care situation in foster care, kinship care, or adoption probation; living in a foster, kinship, or adoption-probation home; assessed by their child welfare worker as likely to remain in their current placement for the duration of the study.
	Educational status enrolled in school grades 1–12
	Language fluent in English (the TYCW program existed only in English);
Exclusion criteria	Special educational needs
	intellectually disabled or very behaviorally disturbed (and thus not likely to complete or benefit from the intervention). Care situation

	<p>living in a group home</p> <p>Education level</p> <p>in the judgment of the child welfare worker, were either very strong students (and thus not likely to need tutoring)</p>
Sample size	83
Split between study groups	<p>15 week TYCW = 42</p> <p>25 week TYCW = 41</p>
Loss to follow-up	<p>15 week TYCW = 5</p> <p>25 week TYCW = 5</p>
% Female	not reported for total sample
Mean age (SD)	not reported for total sample
Outcome measures	<p>Educational outcome 1</p> <p>Woodcock-Johnson-Third Edition (WJ-III; Woodcock, McGrew, & Mather, 2001). The WJ-III is a norm-referenced, standardized test that assesses basic reading and math skills. The following subtests were administered: Letter-Word Identification, Reading Fluency, Story Recall, Understanding Directions, Calculation, Math Fluency, Spelling, Passage Comprehension, Applied Problems, and Story Recall-Delayed, Picture Vocabulary, and Oral Comprehension. These subtests were selected because their administration allows for the calculation of an “intra-achievement” discrepancy score. That is, an Oral Language score (derived from Understanding Directions, Picture Vocabulary, and Oral Comprehension) can be used to predict the level of math and reading achievement, based upon the individual’s level of oral language development. A significant discrepancy between Oral Language ability and academic performance (i.e., reading and math) may help substantiate the existence of a specific math or reading learning disability.</p> <p>Educational outcome 2</p> <p>A Reading Composite score (i.e., Broad Reading) is obtained by combining the Letter-Word Identification, Reading Fluency, and Passage Comprehension subtests.</p> <p>Educational outcome 3</p>

A Math Composite score (i.e., Broad Math) is obtained by combining the scores for Calculation, Math Fluency, and Applied Problems.

Educational outcome 4

Comprehensive Executive Function Inventory—Parent Version (CEFI; Naglieri, 2012). The CEFI is a norm-referenced, standardized measure of executive functioning in children aged 5–18 years. Lower scores indicate greater difficulty with executive functioning.

Educational outcome 5

Parental Support for Learning Scale – Caretaker Version (PSLS; Rogers, Markel, Midgett, Ryan, & Tannock, 2014; Rogers, Hickey, Wiener, Heath, & Noble, 2018). The PSLS was used to assess the extent of caregiver support for educational activities in the home. The PSLS, consisting of 48 items (1 = strongly disagree; 5 = strongly agree), was completed by the caregiver. Two subscales were computed. Instrumental Parental Involvement ($\alpha = 0.87$ in the present sample) assessed the degree of caregiver warmth, patience, and independence regarding the child's school-related choices, with higher scores suggesting more effective involvement. Controlling parental involvement measured caregiver use of commands, punishment, nagging and disapproval regarding the child's schoolwork, with higher scores indicating less effective involvement.

Mental health outcome 1

Strengths and Difficulties Questionnaire. The SDQ uses 25 parent or caregiver ratings to assess mental health problems over the last six months in children or youth aged 4–17 years. The Total Difficulties score was used for the current study, with scores ranging from 0 to 40, with a higher score demonstrating greater behavioural problems.

Mental health outcome 2

Trauma Symptom Checklist for Children (TSCC; Briere, 1996). The TSCC is a self-report instrument that assesses a broad range of symptoms

of traumatic experiences in children and adolescents, aged 8–17 years. Given its reading level, the measure was administered to children aged 10 years and older. Authors used the total Posttraumatic Stress score, based on all 44 items.

Mental health outcome 3

Trauma Symptom Checklist for Young Children (TSCYC; Briere et al., 2001). The TSCYC is a parent-reported measure of traumatic symptoms experienced by young children, ages 3–12 years. In the current study, it was administered to caregivers of foster children aged 5–9 years. A total Posttraumatic Stress T-score was calculated, with a higher T-score indicating a greater level of posttraumatic stress.

Study arms

Teach Your Children Well tutoring (short version) (N = 36)

TYCW program. All of the children in care received the Teach Your Children Well tutoring intervention, for either 15 or 25 weeks. The TYCW program consists of a four-level series of books, written by the designer of the program, Michael Maloney et al. For each reading level, there was a detailed instructor's manual and a student reader and a student workbook. The math program consisted of a four-level series of workbooks and a student workbook. The TYCW tutoring program was designed to provide 3.0 h a week of individual tutoring, that is, two 1.5 h sessions, each divided into 30 min of one-to-one direct instruction in reading, 30 min of one-to-one direct instruction in math, and the remaining 30 min in either math or reading, depending on the needs of the child. The targeted number of TYCW sessions was 30 (i.e., 45 h of tutoring) for participants in the 15-week group.

% Female	44.4%
Mean age (SD)	10.28 ± 2.78 years
Condition specific characteristics	<p>Disabilities, speech or communication needs, or special education needs ADHD: 22.2% Learning Disability: 19.4% Developmental Disability: 16.7% Autistic Spectrum Disorder: 8.1% Psychiatric: 19.4%</p> <p>Care characteristics age of first placement: 4.99 ± 3.14 years Reasons for Entry into care - Neglect: 72.2% Sexual Abuse: 0.0% Domestic Violence: 22.2% Emotional Harm: 13.9% Abandonment: 5.6% Problem Behaviour: 8.3% Other: Parental mental illness: 0.0%</p> <p>Number of placement moves Number of previous placements: 2.30 ± 1.74 Number of unplanned school changes: 1.84 ± 1.50</p>

Teach Your Children Well (long version) (N = 36)

TYCW program. All of the children in care received the Teach Your Children Well tutoring intervention, for either 15 or 25 weeks. The TYCW program consists of a four-level series of books, written by the designer of the program, Michael Maloney et al. For each reading level, there was a detailed instructor's manual and a student reader and a student workbook. The math program consisted of a four-level series of workbooks and a student workbook. The TYCW tutoring program was designed to provide 3.0 h a week of individual tutoring, that is, two 1.5 h sessions, each divided into 30 min of one-to-one direct instruction in reading, 30 min of one-to-one direct instruction in math, and the remaining 30 min in either math or reading, depending on the needs of the child. The targeted number of TYCW sessions was 50 TYCW sessions (or 75 h) for those in the 25-week group.

% Female	44.4%
Mean age (SD)	12.18 ± 2.97 years
Condition specific characteristics	<p>Disabilities, speech or communication needs, or special education needs ADHD: 38.8% Learning Disability: 22.2% Developmental Disability: 11.1% Autistic Spectrum Disorder: 0.0% Psychiatric: 19.4%</p> <p>Care characteristics age of first placement: 5.80 ± 3.86 years Reasons for Entry into care - Neglect: 69.4% Sexual Abuse: 5.6% Domestic Violence: 27.8% Emotional Harm: 38.9% Abandonment: 11.1% Problem Behaviour: 13.9% Other: Parental mental illness: 2.8%</p> <p>Number of placement moves Number of previous placements: 2.61 ± 2.30 Number of unplanned school changes: 1.61 ± 1.41</p>

Risk of Bias

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Some concerns <i>(there were some significant differences observed between comparison groups, slightly more than would be expected by chance. However, these differences were not found to be associated with the outcomes of interest, according to the authors.)</i>
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	High <i>(over 10% drop out in both arms and these results were excluded from the analysis, even where attendance of the intervention had begun)</i>
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low <i>(All of the variables had less than a 6% missing data rate, with the majority having less than 5% missing.)</i>
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Some concerns <i>(Outcome assessors appeared to be unblinded, which may have influenced results)</i>

Section	Question	Answer
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Some concerns (<i>Raw pre-test and post-test data was not presented, by comparison group.</i>)
Overall bias and Directness	Risk of bias judgement	High
	Overall Directness	Indirectly applicable (<i>Canada</i>)

Leve 2007

Study type	Randomised controlled trial (RCT)
Study location	USA
Study setting	Group care and foster care settings
Study dates	1997 to 2002
Duration of follow-up	12 months
Sources of funding	Support for this research was provided by the Oregon Youth Authority and by the following grants: MH54257, NIMH, U.S. PHS; DA15208, NIDA, U.S. PHS; and DA17592, NIDA, U.S. PHS.
Inclusion criteria	<p>Age 13 to 17 years old</p> <p>Care situation Placed in out of home care within 12 months following referral</p> <p>Criminal characteristic Referred by juvenile court judges in Oregon State. At least one criminal referral in the past 12 months</p> <p>Pregnancy Not currently pregnant</p>

	Gender female
Sample size	81
Split between study groups	37 were randomised to MTFC, 44 to Group Care
Loss to follow-up	90% of the sample participated at 3–6 months postbaseline, 88% of the sample participated at 12 months postbaseline, and 12-month lockup data were available for 98% of the sample.
% Female	100%
Mean age (SD)	15.3 ± 1.1 years
Condition specific characteristics	<p>At risk or victims of exploitation 88% had documented physical abuse and 69% had documented sexual abuse</p> <p>Behavior that challenges Prior to entering the study, the average lifetime criminal referrals per girl was 11.9 (SD = 8.9), and 70% of the girls had committed at least one felony</p> <p>Non-white ethnicity 26%</p> <p>Care characteristics At baseline, 68% of the girls had been residing in single-parent families,</p>
Outcome measures	<p>Educational outcome 1 Homework completion: caregivers and girls reported independently at baseline and at 12 months postbaseline on the number of days in the last week that the girls spent at least 30 min/day on homework. In the second measure, caregivers and girls reported on whether or not the girls did homework that day (0 [No]; 1 [Yes]) via three PDR phone interviews conducted within a 1-week period at 3–6 months postbaseline. Scores were aggregated within rater across calls. Composite scores were formed for each of the educational engagement variables by aggregating caregiver and girl reports.</p> <p>Educational outcome 2 School attendance: at 12-months post baseline, caregivers and girls reported of how often the girls attended school (1 [Not attending], 2 [Attending very infrequently], 3 [Attending infrequently], 4 [Attending more often than not], 5 [Attending regularly], or 6 [Attending 100% of the time]). Composite scores were formed for each of the educational engagement variables by aggregating caregiver and girl reports.</p>

Study arms**Multidimensional Treatment Foster Care (MTFC) (N = 37)**

The MTFC model was individualized based on the girls' behavioral problems and on aftercare considerations. The program supervisor placed girls individually in foster homes with trained MTFC foster parents. The program supervisor worked with juvenile justice and school systems and supervised all other MTFC staff involved with the girls and families (e.g., foster parents, skills trainers, and family and individual therapists). Youth behaviors were tracked via the Parent Daily Report Checklist, which is a brief telephone interview conducted each weekday to track foster parents stress level, girl behavior at home and in school, and girl performance on the point-and-level system. Foster parents were trained and supervised to consistently reinforce high rates of positive and normative youth behaviors. When problem behaviors were identified, the program supervisor and foster parents worked to identify a nondegrading definition of the behavior. Typically, the prosocial alternative to the problem behavior was identified (e.g., accepting feedback without comment); once a behavior had been identified and defined for a particular girl, it was included on the point-and-level system that the foster parents implemented at home. The program supervisor coached the foster parents to take points away for all negative behaviors and to give points for all prosocial or adaptive behaviors. An individual therapist met weekly with each girl to focus on problems at school, with her parents, and in the foster home. Targets for the individual therapy sessions were selected based on PDR data, the daily school cards, and the aftercare resources; efforts were then made to motivate the girl to address behaviors that appeared to be having a negative impact. The focus was on adaptive functioning and highlighting the girl's strengths. Thus, each therapist–youth dyad generated mutual definitions of problematic life areas and selected behavioral areas to focus on. Coordinated psychiatric consultation was available when medication management was needed. To help generalize developing skills to environments outside of the foster home, each girl was assigned a skills trainer (typically a recent college graduate), who helped the girl to identify and participate in community activities of interest. The skills trainer also addressed specific social skills by coaching or reinforcing the girl with adaptive ways to respond to specific situations. Once a behavioral target had been identified and clearly defined, the skills trainer attempted to help the girl to expand her behavioral options through role-plays in hypothetical situations and real-world contexts. In many cases, the skills trainer offered to teach appropriate behaviors to prevent the girl from losing points or to help her in earning a desired reinforcer. This approach helped to establish a collaborative relationship. As the skills trainer worked with the youth to develop more adaptive individual behaviors, the family therapist worked with the youth's family to identify prosocial and problem behaviors occurring in the family context and to define structured responses to these behaviors. The family therapist worked with the aftercare resource (typically a biological parent) to improve their supervision, reinforcement, and limit-setting methods. Parents were taught to use the point-and-level system to provide

feedback and consequences for youth behavior using brief, nonemotional reactions to misbehavior, thus avoiding long discussions of the circumstances surrounding the behavior.

% Female	100%
Mean age (SD)	Not reported
Outcome measures	<p>Educational outcome 1 Homework completion score at 3-6 months post-intervention: mean 1.71 ± 1.07; Homework completion score at 12 months post-intervention: mean 3.47 ± 2.44. In multivariable analysis adjusting for baseline homework score, girls receiving MTFC spend significantly longer on homework ($P < 0.01$)</p> <p>Educational outcome 2 School attendance at 12-months post baseline (mean score): 5.48 ± 0.77</p>

Group Care control (N = 44)

Group Care (GC) is the standard intervention service provided for delinquent girls who are referred for out-of-home care. In the current study, girls randomly assigned to the GC condition took part in 1 of 19 community-based group care programs located throughout Oregon State. The programs had 2–51 youth in residence (mean = 21), 1–50 staff members (Median = 2), and on-site schooling. Although each GC program differed somewhat in its theoretical orientations, 86% of the programs endorsed a specific treatment model, of which the primary philosophy of their program was a behavioral (70%), an eclectic (26%), or a family-style therapeutic approach (4%). Seventy percent of the programs reported delivering therapeutic services at least weekly.

% Female	100%
Mean age (SD)	Not reported
Outcome measures	<p>Educational outcome 1 Mean homework completion score at 3-6 months post-baseline: 1.07 ± 1.13; mean homework completion score at 12 months post baseline: 2.03 ± 2.12</p> <p>Educational outcome 2 School attendance mean score at 12-months post baseline: 4.87 ± 1.33</p>

Risk of bias	<p>Domain 1: Bias arising from the randomisation process</p> <p>Some concerns</p> <p>(Unclear how randomisation was performed or if allocation concealment)</p> <p>Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)</p> <p>Some concerns</p> <p>(Unclear if all participants assigned to their groups received their interventions as allocated. Intention to treat analysis used.)</p> <p>Domain 3. Bias due to missing outcome data</p> <p>High</p> <p>(Over 10% lost to follow up. Unclear how much additional missing outcome data or if this differed between comparison groups)</p> <p>Domain 4. Bias in measurement of the outcome</p> <p>Some concerns</p> <p>(Quite crude measures used for homework completion and school attendance. Unclear if outcome assessors were aware of intervention group. Possibility that reporting of outcomes was affected by knowledge of intervention group.)</p> <p>Domain 5. Bias in selection of the reported result</p> <p>Some concerns</p> <p>(In sufficient information to convince that trial was conducted according to a prespecified plan that was finalised before unblinded outcome data was available.)</p> <p>Overall bias and Directness</p> <p>Risk of bias judgement</p> <p>High</p> <p>Overall Directness</p> <p>This question has not yet been answered.</p>
---------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Mooney 2016

Study type	Randomised controlled trial (RCT)
Study location	United Kingdom
Study setting	Children in foster care
Study dates	April 2013 to June 2014
Duration of follow-up	Four weeks following the intervention (which took place over 6 months)
Sources of funding	Funded by UK government. In this period Booktrust received financial support from the Department for Children, Schools and Families to extend its programme to 1600 children. Subsequently, and in 2009, the Letterbox Club opened to every Local Authority in the UK. In the same year The Letterbox Club was introduced as a pilot scheme in Northern Ireland where it has been funded through a partnership between the charities Booktrust and the Fostering Network's Fostering Achievement Scheme since that time.
Inclusion criteria	Age Aged 7-11 years Care situation All children in foster care in Northern Ireland
Sample size	116
Split between study groups	60 children allocated to the letter box intervention and 56 children to the wait list control
Loss to follow-up	4 children were lost to follow up in the control group
% Female	50.9%

Mean age (SD)	Not reported (48.3% aged 7-8 years and 51.7% aged 9-10 years)		
Condition specific characteristics	Care characteristics 67.2% in foster care, 32.8% in kinship care,		
Outcome measures	<p>Educational outcome 1 The Neale Analysis of Reading Ability was used to measure literacy outcomes for the trial (reading rate, accuracy and comprehension)</p> <p>Educational outcome 2 The Elementary Reading Enjoyment Scale (known as the 'Garfield Test') was used to measure the children's attitudes to recreational reading and academic reading.</p> <p>Educational outcome 3 Children were asked "Do you like school?" with the option of reply "not really", "a little" or "a lot".</p> <p>Educational outcome 4 Children were asked "Do you like reading?" with the option of reply "not really", "a little" or "a lot".</p>		
Study arms	<p>Letterbox club (N = 56) The Letterbox Club is a book gifting intervention that provides direct support to children in foster care care aged 7–11 years to improve their educational outcomes. The intervention comprises once-monthly personalised parcels posted between May and October of each year to children in their foster homes. Parcels comprise a brightly coloured envelope (with different colours depending on the age group targeted) which is personally addressed to the child at their foster carers home and which has, as its contents: a personalised letter; two books (one fiction and one non-fiction which have been selected by a panel at Booktrust); stationery items (for example pencils, exercise book, stickers); and a mathematics game (comprising puzzle sheets/practice papers, games with a die/plastic coins for example). The parcels are delivered between May and October each year and over the six-month period it is anticipated that children will have built up their own collection of books and related items consisting of a range of books including non-fiction (biology, history), activity-based, fun based, story based books as well as a book of poems and other items (Winter et al., 2011). As a book gifting scheme directed at the child, the intervention does not rely on, expect or demand foster carer involvement and, as such, there is no manual or guidance for carers about how and in what ways they/the child should engage with the parcel.</p> <table border="1"> <tr> <td>% Female</td> <td>50.0%</td> </tr> </table>	% Female	50.0%
% Female	50.0%		

Mean age (SD)	Not reported (48.2% aged 7-8 years and 51.8% aged 9-10 years)
Condition specific characteristics	Care characteristics 66.1% in foster care, 33.9% in kinship care
Outcome measures	<p>Educational outcome 1 Reading skills mean scores: reading accuracy: 92.15 ± 15.14; reading comprehension: 91.48 ± 16.05; and reading rate: 97.83 ± 14.04</p> <p>Educational outcome 2 Recreational reading mean score: 29.58 ± 7.36; academic reading mean score: 29.78 ± 7.22</p> <p>Educational outcome 3 Liked school "a lot": 55.2%</p> <p>Educational outcome 4 Liked reading "a lot": 58.6%</p>
Wait list (N = 51) The control group did not receive their parcels during the study period but carried on as normal.	
% Female	51.7%
Condition specific characteristics	Care characteristics 68.3% in foster care, 31.7% in kinship care,
Outcome measures	<p>Educational outcome 1 Reading skills mean scores: reading accuracy: 91.15 ± 14.26; reading comprehension: 91.97 ± 15.35; and reading rate: 100.98 ± 13.06</p> <p>Educational outcome 2 Recreational reading mean score: 30.39 ± 6.68; academic reading mean score: 30.45 ± 6.75</p> <p>Educational outcome 3 Liked school "a lot": 65.4%</p> <p>Educational outcome 4 Liked reading "a lot": 61.3%</p>

Risk of bias	Domain 1: Bias arising from the randomisation process
	Low
	Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)
	Low
	(Only participants with missing outcome data excluded and less than 5% attrition)
	Domain 3. Bias due to missing outcome data
	Low
	(<5% missing)
	Domain 4. Bias in measurement of the outcome
	Some concerns
(Possible that tests administered could have been influenced by knowledge of intervention group, but unlikely)	
Domain 5. Bias in selection of the reported result	
Low	
(No protocol cited but enough information provided in the study)	
Overall bias and Directness	
Risk of bias judgement	
Low	
Overall Directness	
This question has not yet been answered.	

Zinn 2014/Courtney 2008

Study type	Randomised controlled trial (RCT)
Study location	USA
Study setting	Home-delivered tutoring for youth in foster care
Study dates	September 2003 to June 2004
Duration of follow-up	Two follow up interviews approximately 13 months apart. the median duration between baseline and second follow up was 26.8 months.
Sources of funding	funding from the Office of Planning, Research and Evaluation and the Children's Bureau, Administration for Children and Families, U.S. Department of Health and Human Services
Inclusion criteria	<p>Age Age 14 or older</p> <p>Care situation In foster care (kinship and non-kinship)</p> <p>Educational status determined by the program to be one to three years behind grade level in reading or math based on an educational assessment process devised by The Community College Foundation</p>
Exclusion criteria	<p>Care situation Not in residential care or correctional placements</p>
Sample size	529
Split between study groups	277 randomised to the intervention group, 252 randomised to the control group

Loss to follow-up	64 youth were considered to be "out of the sampling frame" leading to 31 excluded post-randomisation in the intervention group and, 33 excluded post-randomisation in the control group
% Female	54.4%
Mean age (SD)	14.5 ± 0.8
Condition specific characteristics	<p>Mental health or emotional wellbeing needs PTSD: 6.5%</p> <p>Disabilities, speech or communication needs, or special education needs Participates in special education 35.1%; has learning disability 26.1%</p> <p>Behavior that challenges Internalising behaviours: 30.3%; externalising behaviours: 25.8%; prior runaway from care: 16.6%</p> <p>Non-white ethnicity 69.2%</p> <p>Care characteristics Kinship foster care: 46.5%; Non-kinship foster care: 50.3%; Group home/residential care; other: 0.7%</p>
Outcome measures	<p>Educational outcome 1 Woodcock–Johnson Tests of Achievement III: letter– word identification, calculation, and passage comprehension. These tests provide age-based norms (i.e., percentile scores) for individuals 2 years of age and older.</p> <p>Educational outcome 2 Grade Point Average. Youths were asked what grades they had received in (1) English or language arts, (2) mathematics, (3) history or social studies, and (4) science during their last full semester of school attendance. Response options ranged from "A" (4) to "D or lower" (1). Reported grades in these four subjects were then averaged to obtain an overall grade score. Responses were scored based on a standard 4-point scale, and an overall GPA was computed by taking the average of these.</p> <p>Educational outcome 3 Qualifications: youth were asked several questions about their educational achievement, including the highest grade they had completed, and whether they had a high school diploma or general equivalency diploma</p> <p>Educational outcome 4 School behaviour: School behaviours: Youths were asked to indicate how often they had had "trouble" completing the following five tasks during their last full semester of school attendance: (1) getting along with your teachers, (2) paying attention in school, (3) getting your homework done, (4) getting along with other students, and (5) arriving on time for class. Response options ranged from "never" (0) to "every day" (5). School behaviour was then operationalized or defined as the mean of these five items.</p>

Study arms	<p>ESTEP tutoring program (N = 212)</p> <p>The primary objectives of the ESTEP Tutoring program were to (1) improve the reading and math skills of foster youth, ages 14 and 15, who were one to three years behind grade level in reading or math and (2) empower youth to use other educational services and resources that may have been available to them. At the time of referral to the ESTEP Tutoring program, youth were matched to tutors based on several criteria, including gender congruence and the proximity and availability of the tutor. Tutors were typically undergraduate and graduate students with at least 15 semester credits and grade point averages higher than 2.5 (L.A. DCFS, 2002). The Community College Foundation provided tutors with a one-day training at the start of their employment as well as ongoing training about twice per year. These training sessions cover assessment and curriculum materials, methods to engage youth, and various issues related to case management. After a youth was matched with a tutor, the tutor conducted a series of home visits, during which they assessed the youth's reading, math, and spelling to determine the curriculum levels to use with the youth. The tutor then met individually with the youth twice per week in the youth's home, providing up to 50 hours of tutoring; on average, youth received 8 h of math tutoring and 17 hours of reading (language) tutoring during the evaluation. The curriculum used by ESTEP tutors is based on a combination of the Houghton–Mifflin curricula for math, spelling, and vocabulary, and Science Research Associates Reading 3A curriculum. Generally, tutors reported that they had no involvement with the participating youths' schools, because the ESTEP Tutoring program had an independent curriculum and focus. The ESTEP Tutoring program was a one-to-one, in-home model focused on ameliorating youths' broader academic deficits. Based on program materials and interviews with program staff, the use of a one-to-one model was intended to abet the development of a mentoring relationship between tutors and youth. In addition to facilitating the task of tutoring. This relationship was posited to provide the youth with the skills and experience to develop other healthy relationships with adults. The provision of tutoring in a youths' home was posited to encourage caregivers to become involved and invested in youths' education. Together, these added supports were expected to increase youths' engagement with the ESTEP Tutoring program and, in turn, lead to greater investment by youth in their educational progress.</p> <table border="1" data-bbox="452 1177 2020 1433"> <tr> <td data-bbox="452 1177 689 1433"> Outcome measures </td> <td data-bbox="689 1177 2020 1433"> <p>Educational outcome 1 Woodcock–Johnson Tests of Achievement III (mean scores): letter– word identification: 28.9 ± 22.6; Calculation: 19.5 ± 19.0; passage comprehension 23.0 ± 20.6</p> <p>Educational outcome 2 Grade level completed: 10.2 ± 0.9; Grade point average: 2.3 ± 0.8</p> <p>Educational outcome 3</p> </td> </tr> </table>	Outcome measures	<p>Educational outcome 1 Woodcock–Johnson Tests of Achievement III (mean scores): letter– word identification: 28.9 ± 22.6; Calculation: 19.5 ± 19.0; passage comprehension 23.0 ± 20.6</p> <p>Educational outcome 2 Grade level completed: 10.2 ± 0.9; Grade point average: 2.3 ± 0.8</p> <p>Educational outcome 3</p>
Outcome measures	<p>Educational outcome 1 Woodcock–Johnson Tests of Achievement III (mean scores): letter– word identification: 28.9 ± 22.6; Calculation: 19.5 ± 19.0; passage comprehension 23.0 ± 20.6</p> <p>Educational outcome 2 Grade level completed: 10.2 ± 0.9; Grade point average: 2.3 ± 0.8</p> <p>Educational outcome 3</p>		

	<table border="1"> <tr> <td data-bbox="448 284 689 405"></td> <td data-bbox="689 284 2022 405"> <p>High school diploma or general equivalency diploma: 19 (9.7%)</p> <p>Educational outcome 4 School behaviour score: 1.06 ± 0.70</p> </td> </tr> <tr> <td colspan="2" data-bbox="448 405 2022 539"> <p>No ESTEP tutoring (N = 190) No description of control intervention. Control participants may have accessed tutoring through other sources.</p> </td> </tr> <tr> <td data-bbox="448 539 689 898"> <p>Outcome measures</p> </td> <td data-bbox="689 539 2022 898"> <p>Educational outcome 1 Woodcock–Johnson Tests of Achievement III (mean scores): letter– word identification: 26.8 ± 21.9; Calculation 19.8 ± 20.9; Passage comprehension: 23.2 ± 21.5</p> <p>Educational outcome 2 Grade level completion: 10.2 ± 1.0; Grade Point Average: 2.3 ± 0.7</p> <p>Educational outcome 3 High school diploma or general equivalency diploma: 21 (9.8%)</p> <p>Educational outcome 4 School behaviour score: 1.08 ± 0.75</p> </td> </tr> </table>		<p>High school diploma or general equivalency diploma: 19 (9.7%)</p> <p>Educational outcome 4 School behaviour score: 1.06 ± 0.70</p>	<p>No ESTEP tutoring (N = 190) No description of control intervention. Control participants may have accessed tutoring through other sources.</p>		<p>Outcome measures</p>	<p>Educational outcome 1 Woodcock–Johnson Tests of Achievement III (mean scores): letter– word identification: 26.8 ± 21.9; Calculation 19.8 ± 20.9; Passage comprehension: 23.2 ± 21.5</p> <p>Educational outcome 2 Grade level completion: 10.2 ± 1.0; Grade Point Average: 2.3 ± 0.7</p> <p>Educational outcome 3 High school diploma or general equivalency diploma: 21 (9.8%)</p> <p>Educational outcome 4 School behaviour score: 1.08 ± 0.75</p>
	<p>High school diploma or general equivalency diploma: 19 (9.7%)</p> <p>Educational outcome 4 School behaviour score: 1.06 ± 0.70</p>						
<p>No ESTEP tutoring (N = 190) No description of control intervention. Control participants may have accessed tutoring through other sources.</p>							
<p>Outcome measures</p>	<p>Educational outcome 1 Woodcock–Johnson Tests of Achievement III (mean scores): letter– word identification: 26.8 ± 21.9; Calculation 19.8 ± 20.9; Passage comprehension: 23.2 ± 21.5</p> <p>Educational outcome 2 Grade level completion: 10.2 ± 1.0; Grade Point Average: 2.3 ± 0.7</p> <p>Educational outcome 3 High school diploma or general equivalency diploma: 21 (9.8%)</p> <p>Educational outcome 4 School behaviour score: 1.08 ± 0.75</p>						
<p>Risk of bias</p>	<p>Domain 1: Bias arising from the randomisation process</p> <p>Some concerns</p> <p>(No information about randomisation process or whether allocation was concealed.)</p> <p>Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)</p> <p>High</p> <p>(12% of randomised participants were excluded immediately following randomisation; While intention to treat analysis was used, there was significant deviations from the intended treatment in both groups. 38.2% of those assigned to the E-STEP group did not receive E-STEP services and 12.3% of those in the control group did receive ESTEP services.)</p> <p>Domain 3. Bias due to missing outcome data</p> <p>High</p>						

(Other than the 12% who were excluded immediately following randomisation, there was also <10% who responded to the follow up surveys. The reasons for this are unclear and may be associated with having poorer school outcomes.)

Domain 4. Bias in measurement of the outcome

Some concerns

(Unclear if assessors were blinded to intervention status. It is possible that they may influence some of the outcomes.)

Domain 5. Bias in selection of the reported result

Some concerns

(Insufficient information provided to convince that trial was conducted according to a pre-specified plan)

Overall bias and Directness

Risk of bias judgement

High

(Study authors note that approximately equal proportions of ESTEP and control groups received some form of tutoring (58.4% vs 60.8%))

Overall Directness

This question has not yet been answered.