

youth futures

FOUNDATION

A network meta-analysis of employment and skills interventions designed to assist young people to enter the labour market in high income countries

Technical Report

David Taylor, Rebecca Featherston, Ellie Ott, Jamie Rowland, Becci Newton, Aron Shlonsky

May 2023



Youth Futures Foundation is the What Works Centre for youth employment. We aim to narrow employment gaps for young people facing the greatest challenges by identifying what works and why and investing in evidence generation to improve policy and practice.

The research team would like to acknowledge the contribution of the following individuals for their assistance in the production of scoping notes, identification of relevant literature, data extraction, quality assessment and research support.

From the Centre for Evidence and Implementation: Sarah McKenzie, Caitlin Clymer, Amy Hall, Felicia Lee, Jade Mitchell, Chloe Ang and Adha Lone.

From the Institute for Employment Studies: Alexandra Nancarrow, Daniel Muir, Seemanti Ghosh, Billy Campbell and Beatrice Rosolin.



MONASH
University



Contents

ABOUT THE RESEARCH TEAM	6
CENTRE FOR EVIDENCE AND IMPLEMENTATION	6
MONASH UNIVERSITY	6
INSTITUTE FOR EMPLOYMENT STUDIES.....	6
EXECUTIVE SUMMARY	8
BACKGROUND	8
CONTEXT.....	8
OBJECTIVES	8
METHODOLOGY.....	9
RESULTS	10
DISCUSSION	12
BACKGROUND & CONTEXT	14
WHAT TYPES OF CHALLENGES DO UNEMPLOYED YOUTH FACE?	14
WHAT TYPES OF SUPPORT AVAILABLE TO UNEMPLOYED YOUNG PEOPLE?	15
WHAT IS THE CONTEXT FOR THIS REVIEW?	15
METHODOLOGY	17
OBJECTIVES	17
PROTOCOL REGISTRATION	17
STUDY ELIGIBILITY CRITERIA.....	17
<i>Types of participants</i>	17
<i>Types of interventions</i>	17
<i>Types of comparators</i>	19
<i>Types of outcomes</i>	19
SETTING OF STUDIES	19
<i>Types of studies</i>	20
SEARCH STRATEGY	20
<i>Including studies of related reviews</i>	20
<i>Additional sources</i>	21
<i>References of included studies</i>	22
STUDY ELIGIBILITY CRITERIA.....	22
STUDY SELECTION.....	23
DATA EXTRACTION.....	23
STUDY CONFIDENCE.....	24
DATA ANALYSIS AND SYNTHESIS	25
<i>Measures of treatment effect</i>	25
<i>Outcome selection</i>	26
<i>Unit of analysis issues</i>	27
<i>Dealing with missing data</i>	27
<i>Repeated measures of participants</i>	28

Quantitative synthesis.....	28
Contextualising results	30
Subgroup analysis	31
Sensitivity analysis.....	32
Assessing publication bias.....	32
Assessing network coherence	32
DEVIATIONS FROM THE PROTOCOL.....	33
RESULTS	34
SEARCH RESULTS	34
CHARACTERISTICS OF INCLUDED STUDIES	34
<i>Location of included studies</i>	36
<i>Excluded studies</i>	38
CONFIDENCE IN INCLUDED STUDIES	38
MODEL SELECTION.....	39
EMPLOYMENT STATUS OUTCOMES	40
<i>Structure of network for employment status</i>	41
<i>Assessing the impact of combinations of components on employment status</i>	43
<i>Assessing the impact of each component on employment status</i>	47
<i>Assessing the impact of interactions between components on employment status</i>	48
<i>Identification of a preferred model for employment status</i>	50
EDUCATION OUTCOMES	50
<i>Structure of network for education completion</i>	51
<i>Assessing the impact of combinations of components on education completion</i>	53
<i>Assessing the impact of each component on education completion</i>	56
<i>Assessing the impact of interactions between components on education completion</i>	56
<i>Identification of a preferred model for education completion</i>	58
SUBGROUP ANALYSIS	58
<i>Study confidence</i>	58
<i>Study location</i>	59
<i>Study population</i>	60
SENSITIVITY ANALYSIS	63
ASSESSING PUBLICATION BIAS	64
ASSESSING NETWORK COHERENCE.....	66
DISCUSSION	67
SUMMARY OF KEY FINDINGS	67
<i>What components of the programmes were effective?</i>	67
<i>Were the components more effective for some subgroups?</i>	68
<i>Interactions between components</i>	68
<i>Consideration of residual components</i>	68
RECOMMENDATIONS FOR PRACTICE AND POLICY	69
RECOMMENDATIONS FOR RESEARCH	70
STRENGTHS AND LIMITATIONS OF THE REVIEW FINDINGS	72

<i>Strengths and limitations of CNMA</i>	72
<i>Limitations of a rapid review methodology</i>	72
<i>Potential for limitations in external validity</i>	73
<i>Limitations in considering residual components</i>	73
<i>Limitations in consideration of follow up time</i>	74
STRENGTHS AND LIMITATIONS OF THE AVAILABLE EVIDENCE	74
<i>Strengths of available evidence</i>	74
<i>Limitations of available evidence</i>	74
<i>Limitations in the reporting of results</i>	75
CONCLUSIONS	76
REFERENCES	77
APPENDIX A ADDITIONAL DETAIL ON THE CHARACTERISTICS OF INCLUDED STUDIES	91
APPENDIX B SUPPLEMENTARY INFORMATION ABOUT STUDY METHODOLOGY	106
APPENDIX C SUPPLEMENTARY INFORMATION ABOUT INCLUDED AND EXCLUDED STUDIES	109
APPENDIX D SUPPLEMENTARY NMA RESULTS	118
APPENDIX E ASSESSING NETWORK COHERENCE	122
<i>Evaluating local incoherence</i>	122
<i>Evaluating global incoherence</i>	122

About the research team

Youth Futures Foundation (YFF) commissioned a consortium consisting of the Centre for Evidence and Implementation (CEI), the Institute for Employment Studies (IES) and Monash University to conduct a series of rapid evidence assessments to inform the development of a Youth Employment toolkit.

Centre for Evidence and Implementation

The Centre for Evidence and Implementation (CEI) is a global, not-for-profit evidence intermediary dedicated to using the best evidence in practice and policy to improve the lives of children, families and communities facing adversity. Established in Australia in late 2015, CEI is a multi-disciplinary team across five offices in London, Oslo, Singapore, Melbourne and Sydney.

They work with their clients, including policymakers, governments, practitioners, programme providers, organization leaders, philanthropists and funders in three key areas of work:

- Understand the evidence base
- Develop methods and processes to put the evidence into practice
- Trial, test and evaluate policies and programmes to drive more effective decisions and deliver better outcomes

Monash University

Monash University, the largest university in Australia, is ranked in the world's top 100 and is a member of the prestigious Group of Eight Australian universities. It is widely recognised as one of the most international universities globally.

The School of Primary and Allied Health Care is part the Faculty of Medicine, Nursing and Health Sciences, one of the world's top health education institutions. Professor Aron Shlonsky, Head of Department – Social Work, leads a team of analysts and methodologists that specialise in applied social research.

They have expertise in experimental and quasi-experimental design; systematic reviews; policy analysis; measuring and accounting for implementation processes and outcomes; large scale data analytics; and the design and use of administrative, survey, and interview data in research.

Institute for Employment Studies

The Institute for Employment Studies (IES) is an independent, apolitical, international centre of research and consultancy in public employment policy and organisational human resource management. It works closely with employers, government departments, agencies, and professional and employee bodies.

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

For over 50 years IES has been a focus of knowledge and practical experience in employment and training policy, the operation of labour markets, and human resource planning and development.

Executive Summary

Background

Youth unemployment is a major challenge worldwide with an estimated 75.1 million young people unemployed in 2021 (International Labour Organization, 2022) and nearly half a million young people unemployed in the UK at the start of 2023 (Buchanan, 2023). Unemployment has an adverse effect on human capital and negatively affects health, happiness, crime levels and socio-political stability (Kluve et al., 2017).

A broad range of policies, programmes and interventions have been implemented to assist young people to develop the skills and gain the experience required to enter the labour market. Many fall under the broad category of Active Labour Market Policies (ALMP), which include investment in public employment services and administration, labour market training and programs that support the transition from school to work as well as supporting employment incentives, entrepreneurship programmes and rehabilitation and desistance programmes (OECD, 2021; White & Apunyo, 2021). Many of these interventions are delivered as part of larger programmes that contain two or more intervention components.

Context

To better understand the impact of such programmes for application in England, Youth Futures Foundation (YFF) commissioned the Centre for Evidence and Implementation (CEI), the Institute for Employment Studies (IES) and Monash University to conduct two rapid evidence assessments (REAs) to inform the development of a Youth Employment toolkit. This REA examines six types of interventions to support youth employment; the other reviews the literature on wage subsidies to employers.¹

Objectives

The objective of this rapid evidence assessment (REA) was to assess the effectiveness of some common constituent components of employment and skills programmes designed to assist young people to enter the labour market in high income countries.

The primary research question that guided this review was:

What combination of components should an employment and skills programme have in order to be effective at supporting young people to enter paid employment?

A secondary research question explored:

¹ The toolkit will provide policymakers and practitioners with information about evidence-informed programmes across a range of programme areas. In its initial form, it will look at seven topic areas; the six interventions examined here, and wage subsidies to employers.

What combination of components should an employment and skills programme have in order to be effective at supporting young people to complete educational qualifications?

Methodology

Search strategy

The REA followed a pre-established protocol, which is available on Open Science Framework at <https://osf.io/8w79s> (Ott et al., 2022).

Published and grey literature were considered eligible for this review. The review adopted a pragmatic approach to identifying relevant studies. The included studies were first identified by screening literature included in an Evidence and Gap Map (EGM) by White and Apunyo (2021). This was supplemented with searches of clearinghouses and organisations known to be undertaking or consolidating research on this topic.

Selection criteria and screening

The population of interest was young people and emerging adults (16-30 years old) who were not currently in formal paid employment. Interventions of interest to this review were employment and skills programmes that included one or more of the following components: apprenticeships, basic skills training, life skills training, on-the-job training, off-the-job training or coaching and mentoring. Comparators included: services as usual, another intervention, no intervention, or wait-list control. The primary outcome was entry to employment post-intervention (hereafter referred to as 'employment status'). The secondary outcome was completion of educational qualifications (hereafter referred to as 'education completion'). Studies used experimental or non-randomised, quasi-experimental designs and were conducted in high-income countries, as defined by the World Bank (2022). The latter was to maximise applicability to the English context.

The title and abstracts of potentially relevant studies identified by the search strategy were reviewed against the inclusion criteria by two reviewers working independently. Any discrepancies were resolved by a third reviewer. The screening process was replicated for potentially relevant full-text articles.

Data collection and analysis

Data were extracted by a single reviewer, with a second expert reviewer overseeing the process and checking data.

Data from the included studies were quantitatively synthesised using a network meta-analysis (NMA) approach. NMA is a statistical technique that can be used to quantitatively synthesise the results from multiple studies that aim to achieve similar outcomes by combining direct and indirect evidence in a network (Tsokani et al., 2022). Most of the studies included in this review evaluated programmes that consisted of multiple components (i.e., they were multi-component interventions). Because of this, we reported the results in different forms. Firstly, we undertook a standard NMA to examine the

impact of combinations of components, i.e., as they were delivered as part of interventions. Secondly, we used a *component* network meta-analysis (CNMA) to disentangle the effect of each component integrated into the larger programs that were evaluated by the studies. The CNMA approach allowed for subgroup analyses, which can be used to determine whether certain well-described populations within samples (e.g., age groups) do better than others when they receive certain components.

Results

Included studies

Sixty (n=60) studies — reported in 73 publications — were included in this review. Thirty-two (n=32) used a randomised study design and the remaining twenty-eight (n=28) used a non-randomised study design.

Of the sixty included studies, two-thirds (n=40) were conducted in the United States. Of the remaining third (n=20), all bar one — which was from Australia — were from Europe.

Study confidence

The confidence that we can have in the study findings (i.e., study quality) was assessed using the *Quality assessment of Impact Evaluations* tool (White et al., 2022).² A majority (n=33; 55 per cent) of the included studies received an assessment of 'low confidence', with the remaining assessed as 'medium confidence' (n=16; 27 per cent) or 'high confidence' (n=11; 18 per cent). Among the 'low confidence' studies, the domains that drove the lower confidence in the study findings were high attrition³ (n=19) and baseline balance⁴ (n=18).

Results of intervention-level quantitative synthesis

The results of the standard NMA — which assessed the effect of combinations of components, i.e., as they are delivered as part of interventions — show that some combinations of components had a positive and statistically significant effect on employment status, but there were none that did the same for education completion.

For employment status, there are five combinations of components that, when delivered together, all show a statistically significant and high impact on employment status relative to services as usual. Ordered from largest to smallest they are:

² The tool scores studies as either low, medium or high confidence across six domains including controlling for confounders, use of adequate sample size, loss to follow up (attrition), intervention description, definition of outcome measures and baseline balance between treatment and comparison groups.

³ If losses to follow up are both presented and acceptable

⁴ If treatment and comparison groups are similar (i.e., balanced across important variables) prior to intervention commencement.

- Two (n=2) studies included in the network combined On-the-job training + Other (g=0.48, 95% CI: [0.11, 0.84], p < 0.01). This combination has an NNT of 5.3 (95% CI: 3.2-22.9) indicating that, on average, for every five individuals who received the intervention one additional individual will subsequently be employed.
- Five (n=5) studies included in the network combined Basic Skills + Off-the-job training + Other (g=0.30, 95% CI: [0.12, 0.48], p < 0.01). The NNT for this combination is 8.4 (95% CI: 5.3-20.9), meaning that, on average, for every eight individuals who received the intervention, one additional individual will subsequently be employed.
- Three (n=3) studies included in the network only included On-the-job training (g=0.25, 95% CI: [0.05, 0.46], p < 0.01). The NNT for this combination is 10.1 (95% CI: 5.5-50.4), indicating that, on average, for every ten individuals who receive the interventions, one additional individual will be employed.
- Six (n=6) studies included in the network combined Life Skills + Coaching & Mentoring + Other (g=0.24, 95% CI: [0.08, 0.39], p < 0.01). With an NNT of 10.5 (95% CI: 6.5-31.5) means that, on average, for every ten individuals who receive the interventions, one additional individual will be employed with this combination.
- Four (n=4) studies included in the network only included Off-the-job training (g=0.23, 95% CI: [0.06, 0.40], p < 0.01). The NNT for this combination is 10.9 (95% CI: 6.3-42), meaning that, on average, for every ten individuals who receive the interventions, one additional individual will be employed.

Results of component-level quantitative synthesis

Several headline results emerged from the additive component⁵ network meta-analysis (CNMA). By outcome, these were:

Employment status

- Off-the-job training (g=0.13, 95% CI: [0.01; 0.25], p < 0.05) had statistically significant, moderate sized impact — meaning those who received off-the-job training were more likely to attain employment than those who received services as usual. To put this in context, this means that for every 19.3 (95% CI: 10.1- 252.5) individuals who receive this component, one additional person will be employed.
- The effects of both apprenticeships (g=0.22, 95% CI: [-0.08; 0.52], p > 0.05) and on-the-job-training (g=0.18, 95% CI: [-0.00; 0.35], p > 0.05) are not statistically significant, however we report them here because there are some indications that the network may be under-powered for detecting small, but meaningful differences, and findings may translate into

⁵ See methodology section for a description of additive component NMA.

significant effects in future analyses as the number of studies with similar findings increase.

- All other components — Coaching and Mentoring, Life Skills, Basic Skills and Other — showed 'low impact' effect sizes that were not statistically significant.
- There were no statistically significant differences between other components of interest to this review — i.e., apprenticeships, on-the-job training, coaching and mentoring, basic skills or life skills — relative to services as usual.
- No statistically significant adverse effects on employment status were identified.

Education completion

- There were no statistically significant differences in education completion between those individuals in included studies that received any of the components of interest to this review (on-the-job training, off-the-job training, coaching and mentoring, basic skills or life skills) relative to services as usual. No statistically significant adverse effects on education completion were identified.

Interactions between components

When examining whether any interactions existed between any two components that were delivered together — that is, whether they had an effect greater or less than the sum of their parts — no statistically significant interactions between the program components were found. This was the case for the analyses conducted for both outcomes.

Subgroup and sensitivity analysis

Subgroup analyses were conducted that considered: study confidence, location of included studies and population needs. A sensitivity analysis was applied to study design. The results revealed that the effect of both on-the-job training ($g=1.6$, 95% CI: [0.90; 2.30], $p < 0.01$) and off-the-job training ($g=0.6$, 95% CI: [0.08; 1.12], $p < 0.05$) on employment status was significantly larger when provided to populations who reported additional barriers (i.e., those living with a disability or with known elevated risks — see Methodology section for definition). One interpretation of this is that young people without additional barriers may be more likely to find employment, while those facing additional barriers appear to benefit from the assistance of these particular components.

Discussion

Recommendations for practice and policy

Young people not in employment, education or training face a range of barriers to securing and maintaining employment. The findings from this

review suggest there is no panacea for this. The major recommendation for practice and policy that emerges from this work relates directly to the key finding from this review — that on-the-job training and off-the-job training have very large effects on employment status for young people who report facing additional barriers. There may be merit in an approach to commissioning that involves the targeted implementation of some programme components to young people facing such challenges.

Recommendations for research

A majority of the included studies were undertaken in the United States. Given that the context and policy setting can vary widely between countries, it is unlikely that study findings are partially or wholly generalisable to other settings. Therefore, this research highlights a clear need for more rigorous primary research on the impact of employment and skills programmes in settings outside the United States. To account for exposure at baseline (i.e., if an individual was employed when the intervention commenced), it would be helpful if future primary research utilised analytic methods that measured the impact of the programme on outcomes over time, i.e., to see if the results between both groups changed over time. Future primary research should also include more detailed information on the study population, which would allow for more in-depth analyses on how intervention effectiveness differs for diverse populations.

Many studies did not comprehensively report details about the interventions or their respective components. Improving reporting on both the content of programmes or interventions (what exactly do they do, and for how long and how intensely do they do it) and what services as usual look like in the setting where the programme is being implemented would be beneficial.

Conclusions

This review found that some components of employment and skills programmes — namely off-the-job training — can have a moderate impact on improving employment outcomes for young people who are not in employment, education or training. The effect of some programme components is amplified considerably when they are provided to young people who report facing additional barriers, suggesting that there is benefit in targeting these components to particular populations.

There are numerous opportunities for future research to strengthen the evidence base, particularly by undertaking primary research outside the United States. There are also opportunities to repeat and extend the methods used in this review to provide additional insights on the impact of other components of employment and skills programmes.

Background & Context

What types of challenges do unemployed youth face?

Youth unemployment is a major challenge worldwide with 75.1 million young people unemployed in 2021 (International Labour Organization, 2022) and nearly half a million young people unemployed in the UK at the start of 2023 (Buchanan, 2023). Unemployment has an adverse effect on human capital and negatively affects health, happiness, crime levels and socio-political stability (Kluve et al., 2017).

While youth unemployment rates are recovering in the UK following the Covid-19 pandemic, young people remain disproportionately impacted. Relative to the older population, young people are more likely to have lost work or to have been unable to enter the labour market during the pandemic (White & Apunyo, 2021). The youth unemployment rate — the proportion of economically active 16 to 24 year olds who are unemployed — stands at 10.8 per cent across the UK at the time of writing (Office for National Statistics, 2023). Youth unemployment rates differ significantly across the UK, ranging from 5.5 per cent in South West England to 15 per cent in London (Office for National Statistics, 2023), suggesting that the challenges faced by young people vary across the country. The UK youth unemployment rate is lower than the European Union average, but is 4 per cent higher than in comparable countries such as Germany, demonstrating the ongoing challenges faced by young people (Buchanan, 2023).

Young people from disadvantaged backgrounds, and from some ethnic backgrounds (such as the British Black, Bangladeshi, and Pakistani communities), are overrepresented among the unemployed and economically inactive in the UK, and this disparity continues into later adulthood (Li & Heath, 2020; Longhi, 2020). This trend has been exacerbated by the impact of Covid-19 in the UK (Learning and Work Institute, 2022).

Likewise, young people with disabilities, including mental health conditions, are less likely than their peers to be employed (Department for Work & Pensions, 2023). Over recent decades, there has been a decrease in economic inactivity among young mothers (as the percentage of parents who are young decreases) and an increase in the economic inactivity of young people due to health problems, with the sharpest increase being economic inactivity due to mental health problems (Murphy, 2022). There are concerns that young people with mental health problems who are not in work or study stay economically inactive for longer, with long-term impacts for individuals and society (Murphy, 2022). Finally, youth unemployment represents a significant opportunity cost for society as the productive potential of these young people is underutilised (White & Apunyo, 2021).

What types of support available to unemployed young people?

A broad range of policies, programmes and interventions have been implemented to assist young people to develop the skills and gain the experience required to enter the labour market. Many fall under the broad category of Active Labour Market Policies (ALMPs), which includes spending on public employment services and administration, vocational training, designing programs for youth when in transition from school to work, employment incentives, start-up support and rehabilitation (OECD, 2021; White & Apunyo, 2021). Many approaches to improving youth employment rates also include personal support such as teaching life skills or providing mentoring or coaching as standalone or supplemental support. Young people with more complex needs such as mental health issues may require more targeted employment support (Murphy, 2022).

What is the context for this review?

Commissioning process

The Youth Futures Foundation (YFF) commissioned the Centre for Evidence and Implementation (CEI), the Institute for Employment Studies (IES) and Monash University to conduct a series of rapid evidence assessments to inform the development of a Youth Employment toolkit.⁶

Establishment of review scope

YFF previously commissioned the Campbell Collaboration to produce an evidence and gap map (EGM) on interventions that improve youth employment outcomes (White & Apunyo, 2021). The review team used the findings of the EGM to determine which types of interventions could be suitable for meta-analysis.⁷ A number of intervention types were defined and mapped for priority and ten scoping notes were delivered to discuss the merits of synthesising each one. In the process of reviewing the literature it was observed that a number of interventions are commonly delivered together. We then decided which studies could be combined in a meta-analysis based on the following criteria:

- if they could be delivered together, or it made sense to consider them together,
- studies examining the intervention existed in sufficient number,

⁶ The toolkit will provide policymakers and practitioners with information about evidence-informed programmes across a range of programme areas. In its initial form, it will look at seven topic areas.

⁷ To determine which topics were suitable to synthesise CEI and YFF, produced a series of ten scoping notes to explore the scope and scale of the literature available on each of the topics of interest to YFF and help refine the initial scope of the toolkit. Scoping notes looked at the following topics: apprenticeships, basic skills, career guidance, employment services, life skills, minimum wage, supported employment, technical and vocational training, employer subsidies and mentoring & coaching.

- the intervention was of interest to the funder and their stakeholders in their consultations.

From this process, seven interventions were identified as being candidates for a rapid evidence assessment, including:

- Subsidies to employers,
- Apprenticeships,
- On-the-job training (i.e., traineeships and structured internships),
- Off-the-job training (i.e., technical and vocational training that is predominantly classroom-based rather than in the employment context),
- Basic skills training (i.e., focused on literacy, numeracy and digital skills),
- Life skills training (i.e., covering the range of soft skills, employability attributes and equipping young people with the life management skills that underpin employment), and
- Coaching and mentoring.

Subsidies to employers was selected for a stand-alone rapid evidence review based on findings from the scoping notes. The other interventions were included in this review.

Challenges in synthesising programmes with multiple components

In the production of the scoping notes, the review team observed that these interventions were often delivered as part of larger programmes that contained two or more components.

This issue was also identified in a systematic review by Kluve et al. (2017), which noted that “...a youth employment programme was considered to be a single entity that might consist of one or several interventions. In addition, each of these interventions could have different components: It was possible to find a comprehensive intervention that offered, for instance, both skills training and employment services (to the same participant).”

Selection of network meta-analysis methodology

The review team identified that a CNMA might be able to disentangle the relative impact of each of the programme components of interest to YFF. A scoping exercise was undertaken to test the feasibility of this idea, and an extensive scoping note reported on this process. Following the scoping exercise, the review team and YFF jointly concluded it was likely to be feasible and a protocol was produced to guide its production.

Methodology

Objectives

The objective of this rapid evidence assessment (REA) is to assess the effectiveness of six common constituent components of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries.

The primary research question that guided this review was:

What combination of components should an employment and skills programme have in order to be effective at supporting young people to enter paid employment?

A secondary research question explored:

What combination of components should an employment and skills programme have in order to be effective at supporting young people to complete educational qualifications?

Protocol registration

The REA followed an explicit protocol. The protocol was reviewed by YFF and its external advisors and is available on Open Science Framework at <https://osf.io/8w79s> (Ott et al., 2022).

Study eligibility criteria

Types of participants

Young people and emerging adults (16-30 years old) who are not currently in formal paid employment.

Types of interventions

Interventions of interest to this review were selected from those identified by the review team during the scoping phase. Programmes or interventions were included if they involved the provision of one or more of the following mutually exclusive components — see Table 1 for definitions:

- Apprenticeships,
- Basic skills training,
- Life skills training,
- On-the-job training,
- Off-the-job training,
- Coaching and mentoring.

Table 1 How we defined included intervention components

TERM	DEFINITION
Apprenticeships	<p>To be considered as an apprenticeship, the intervention component (as described) needs to include:</p> <ul style="list-style-type: none"> • the attainment of skills required for mastery of an occupational skill • both on-the-job training and off-the-job training elements • the on-the-job training component needs to be paid • the off-the-job training component needs to be provided by an accredited learning provider • completion leads to a recognised qualification (either national or state-level) • be at least 12 months in length (distinguishing it from an on-the-job training).
On-The-Job Training	<p>To be considered as 'on-the-job training', the intervention component (as described) needs to include:</p> <ul style="list-style-type: none"> • a formal arrangement between an employer and training/intervention provider where the participant undertakes training on-the-job that leads to the development of practical skills (distinguishing it from basic skills) • a training period that lasts at least six weeks, but is less than twelve months (distinguishing it from an apprenticeship)
Off-The-Job Training	<p>To be considered as 'off-the-job training', the intervention component (as described) needs to include:</p> <ul style="list-style-type: none"> • classroom (or equivalent) based curricula that leads to the development of practical skills (distinguishing it from basic or life skills) • it could contribute toward the achievement of a certificate or qualification (but not a high school or equivalent qualification) • will typically last at least 6-12 months • training is provided on a full-time basis
Basic Skills Training	<p>To be considered as 'basic skills' training, the intervention component (as described) needs to include training in a fundamental skill that is essential for re-engaging with education or attaining employment. These could include things such as literacy and numeracy and digital skills. Note, training does not need to lead to any formal qualification.</p>
Life Skills Training	<p>To be considered as 'life skills' training, the intervention (as described) needs to include:</p> <ul style="list-style-type: none"> • Training in 'soft skills' that help communicate and build relationships, emotional intelligence, confidence etc., or • Training in basic practical skills for day-to-day life such as self-care and financial literacy.
Coaching & Mentoring	<p>To be considered as 'coaching and mentoring', the intervention (as described) needs to include a structured mentoring or coaching component that is a formal part of the programme or intervention.</p>

Types of comparators

The following comparisons were included:

- intervention compared with services as usual (i.e., what an individual would have received if they did not receive the intervention),
- intervention compared with another intervention (i.e., another employment and skills programme),
- intervention compared with no intervention (i.e., similar to services as usual, expect there are no alternative services), or
- intervention compared to wait-list control (i.e., comparison group is drawn from waiting list for intervention).

Types of outcomes

Outcomes were considered if they were obtained by analysis of administrative data, survey or interview.

Primary outcomes

The primary outcome was entry to employment post intervention. The review team considered any outcome that represented an individual's subsequent employment status such as:

- Employment status
- Hours worked
- Earnings and salary

Secondary outcomes

The secondary outcome was completion of educational qualifications. The review team considered any outcome that represented an individual's completion of an educational qualification. These were:

- Secondary school, high school or equivalent completion
- Vocational education commencement
- University commencement

Setting of studies

The review included studies conducted in educational, employment, or community settings (e.g., delivered by non-government/third sector organisations or local government authorities).

Studies needed to be conducted in high-income countries, as defined by the World Bank (2022).

Types of studies

The following experimental and quasi-experimental study designs were included:

- Randomised Controlled Trials (RCTs) – including individual RCTs, cluster RCTs and Step-Wedge designs with random time allocation.
- Non-randomised studies that use quasi-experimental methods – including difference-in-difference estimation, synthetic control group methods, studies based on covariate matching, propensity score-based methods, doubly robust methods, regression adjustment, regression discontinuity designs, instrumental variable estimation and non-equivalent control group designs using parallel cohorts that adjust for baseline equivalence.

Search strategy

Including studies of related reviews

This review adopted a pragmatic approach to identifying relevant studies.

Multiple evidence synthesis products have explored this topic area in the last five years. We sought to leverage this work by screening literature identified in these pieces of work.

An Evidence and Gap Map (EGM) by White and Apunyo (2021) already integrated a number of resources including an EGM by the International Initiative for Impact Evaluation (2017) titled *Youth Employment Evidence Gap Map*, and a Campbell Collaboration review by Kluve et al. (2017) titled *Interventions to improve the labour market outcomes of youth: A systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions*. Additionally, we made enquiries throughout the process about an in-progress update to the Kluve et al. review, but it was not available within the relevant timeframe.

All low, medium and high-quality impact evaluations from the White and Apunyo (2021) EGM were screened for relevance using the following intervention categories (as defined by the EGM):

- Life skills
- Internship and apprenticeships
- Employee mentoring (inc. on-the-job training)

Included studies were additionally filtered in accordance with the following outcome categories (as defined by the EGM):

- Employment status and duration
- Hours worked
- Earnings and salary
- Education completion and qualification

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

- Access to/in education
- Technical skills & vocational training

Additional sources

Clearinghouses, government agencies and organisations known to be undertaking or consolidating research in this area were also reviewed using methods detailed in Table 2.

Table 2 Additional sources and methods used to identify relevant studies

SOURCE	SEARCH METHOD	DATE SEARCHED
Pathways to Work Evidence Clearinghouse (United States Administration for Children & Families, 2022)	All programmes and/or interventions for "young adults (aged 16-24)" were reviewed	23/08/2022
Clearinghouse for Labor Evaluation and Research (CLEAR) (United States Department of Labor, 2022)	We searched the term "youth employment" and "evaluation" and screened all studies for relevance.	11/07/2022
United Kingdom Department for Work and Pensions (DWP)	We searched "youth employment" and "evaluation." Additionally, we searched the data archive by category of subject "young people."	11/07/2022
United Kingdom Department for Education (DfE)	We searched the national archives using the adult learning and workplace training filters.	11/07/2022
United States Administration for Children and Families, Office of Planning, Research and Evaluation (OPRE)	We searched using the terms "youth employment" and "evaluat*"	11/07/2022
Organisation for Economic Co-operation and Development (OECD)	We searched using: <i>(All Fields contains "youth employment") from (Language contains 'en') AND from (All Fields contains 'evaluat*') AND from (IGO collection contains "igo/oecd") with type(s) subtype/article OR subtype/workingpaper</i>	11/07/2022
World Bank	We searched using the terms "youth employment" and "evaluat*".	11/07/2022
Institute of Labor Economics (IZA)	We searched using the terms "youth employment" and "evaluat*"	11/07/2022
MDRC	We searched using the term "youth employment" with the publication	11/07/2022

SOURCE	SEARCH METHOD	DATE SEARCHED
	filter set to report/working paper/brief.	
Google Scholar	The first fifty results using the following search terms were reviewed: "Youth Employment" + "evaluat*	11/07/2022

References of included studies

References from all included studies were also screened for inclusion.

Study eligibility criteria

Inclusion and exclusion criteria organised by PICOSS domain are detailed in Table 3 below.

Table 3 Inclusion and exclusion criteria

PICOSS	INCLUSION CRITERIA	EXCLUSION CRITERIA
Population	Young people aged between 16 and 30.	Young people aged less than 16 or over 30.
Intervention	Programmes or Interventions need to involve the provision of one or more of the following components: <ul style="list-style-type: none"> • Apprenticeships, • Basic Skills, • Life Skills, • On-the-job training, • Off-the-job training, or • Coaching and Mentoring — see Table 1 for definitions. 	Interventions that solely involve other components.
Comparison	Usual services, no intervention, other services, or wait-list control.	Studies using other comparators.
Outcome	Studies that examine: <p><i>Primary outcome</i></p> <ul style="list-style-type: none"> • Employment (i.e., employment status, hours worked or earnings and salary) <p><i>Secondary outcome</i></p> <ul style="list-style-type: none"> • Education (i.e., education completion and qualification or access to / in education) 	Studies that examine other outcomes

PICOSS	INCLUSION CRITERIA	EXCLUSION CRITERIA
Study design	<p><i>Experimental designs:</i></p> <ul style="list-style-type: none"> • Randomised Controlled Trials (RCT) including individual RCTs and cluster RCTs • Step-Wedge designs with random time allocation <p><i>Quasi-experimental designs:</i></p> <ul style="list-style-type: none"> • Non-equivalent control group designs using parallel cohorts that adjust for baseline equivalence • Difference-in-Difference estimation • Interrupted time-series • Synthetic control group methods • Studies based on: <ul style="list-style-type: none"> - covariate matching - propensity score-based methods, - doubly robust methods - regression adjustment - regression discontinuity designs, and - instrumental variable estimation 	<p>Non-primary studies, including:</p> <ul style="list-style-type: none"> • Literature reviews • Systematic reviews • Meta-analyses <p>Studies without a valid counterfactual, including designs that do not use a parallel cohort that establishes or adjusts for baseline equivalence, including:</p> <ul style="list-style-type: none"> • Single group pre-post designs • Control group designs without matching in time and establishing baseline equivalence • Cross-sectional designs • Non-controlled observational (cohort) designs • Case-control designs • Case studies / series • Surveys
Setting	Studies undertaken in high income countries as defined by the World Bank (2022).	Studies undertaken in low or middle income countries as defined by the World Bank (2022).
Other	Studies published in English.	Studies published in languages other than English.

Study selection

Potential studies were identified from relevant existing reviews. Citations were moved into Mendeley reference manager for deduplication, subsequently uploaded to Covidence (systematic review software) for screening. Title and abstracts were reviewed against the inclusion criteria by two reviewers working independently, with any discrepancies resolved by a third reviewer. The full texts of any potentially relevant studies were then screened again by a further two reviewers working independently, with any conflicts resolved by a third reviewer.

Data extraction

The data from included studies were extracted by a single reviewer — with a second experienced reviewer overseeing the process and checking the

data. Queries were also raised in weekly team meetings and regular communications. Authors were contacted for missing data. Table 4 summarises the information that was extracted at this stage.

Table 4 Information extracted from included studies

CATEGORY	DETAIL
Bibliographic information	Author, title, year of publication
Intervention details	Intervention name, relevant components (treatment), relevant components (comparison), year intervention commenced, year intervention ended, year study commenced, year study ended, inclusion criteria, exclusion criteria, intervention length, intervention delivery mode, comparison type, comparison intervention (if relevant)
Setting	Study location (country)
Population	Sample size, # treatment, # comparison, gender (% female), ethnicity, self-reported disability, elevated risk
Study design	Study design, study method
Outcomes	Outcome domain, outcome measure, time of measurement
Results	Result type, reported result (treatment), reported results (comparison), reported standard error (SE) (treatment), reported SE (comparison), reported or derived standard deviation (SD) (treatment effect or treatment), reported or derived SD (comparison), reported treatment effect (TE) type, reported TE, reported TE SE, reported TE 95 per cent confidence interval (CI) (lower), reported TE 95 per cent CI (upper), reported TE t-stat, stat. sig (p-value), reported effect size (ES) type, reported ES, reported ES 95 per cent CI (lower), reported ES 95 per cent CI (upper), reported ES t-stat, ES stat. sig (p-value)

Study confidence

Confidence in included studies was assessed using the *Quality assessment of Impact Evaluations* tool (White et al., 2022), in alignment with the EGM on *The effectiveness of Interventions to improve employment*. Study confidence was assessed by one reviewer, with the results checked by another.

The tool scores studies as either low, medium or high confidence across six domains:

1. If the study design can control for potential confounders
2. If the study has adequate sample size
3. If losses to follow up are presented and acceptable
4. If the intervention is clearly defined
5. If outcome measures are clearly defined
6. If there is baseline balance between treatment and comparison groups

Overall confidence in study findings is calculated by taking the lowest rating across domains 1, 3, 5 and 6.

Data analysis and synthesis

Measures of treatment effect

Selecting from multiple reported results

Some studies reported multiple treatment effects from different regression model specifications. The review team developed and applied the following hierarchy to assist in the selection of model results:

- Intention to Treat (ITT)
- Average Treatment Effect (ATE)
- Local Average Treatment Effect (LATE), a.k.a. Complier average causal effect (CACE)
- Average Treatment Effect on the Treated (ATET), a.k.a. Treatment on the Treated (TOT)

In addition to this selection hierarchy, where authors reported both means and regression adjusted means, regression adjusted means were used.

Selecting a common effect size

Studies reported quantitative results in a range of forms, some with effect sizes, and some without. Based on the types of reported results, the Standardised Mean Difference (SMD) was selected as the most appropriate effect size to use for our synthesis. This judgement was influenced by two factors: a) a wide range of results can be easily transformed to an SMD and b) if the review was able to transform continuous outcomes (i.e., hours worked and earnings/wages) then it would allow for a common comparator.

For every included study, effect sizes needed to be estimated from available data, while transformation was required in others.

Pre-transformations required for estimating effect sizes

For some studies, additional transformations were required to obtain information to calculate an effect size. These included:

- Where results were only reported at the subgroup level, combining results to get a result for the whole treatment group — e.g., combining Male and Female study participants.
- Where they were not reported, estimating treatment and comparison sample sizes by using the assignment proportion reported by the authors.
- Where outcomes were only reported graphically, using plot digitizer software to extract estimates (*PlotDigitizer Online App*, n.d.).

A catalogue of instances where we undertook these actions is included in Table 9 in Appendix A.

Processes for transforming effect sizes

Effect sizes were transformed using the 'esc' package developed by Lüdtke (2019) — which is an R implementation of Wilson's (n.d.) Effect Size calculator — for the R Project for Statistical Computing (R Core Team, 2020). Methods used to transform each type of reported result included:

- *Unstandardised regression coefficients* — transformed into SMD (Hedge's *g*) using `esc_B` function
- *Standardised regression coefficients* — transformed into SMD (Hedge's *g*) using `esc_beta` function
- *Count or per cent in each group* — transformed into SMD (Hedge's *g*) using `esc_bin_prop` function
- *Odds ratio* — transformed into SMD (Hedge's *g*) using `or` function
- *Chi-square* — transformed into SMD (Hedge's *g*) using `esc_chisq` function
- *F-stat* — transformed into SMD (Hedge's *g*) using `esc_f` function
- *T-stat* — transformed into SMD (Hedge's *g*) using `esc_t` function

Outcome selection

Employment status

Some studies reported multiple outcomes that investigated the same construct. This was most notably an issue relating to studies reporting different measures of employment status. To select the most appropriate outcome the review team developed a selection hierarchy in cases where multiple outcomes were reported:

- *Ever worked* — an individual was employed at any point, for any duration, after commencement of the intervention,
- *Worked in previous period* — an individual was employed at any point, in a defined period of time prior to measurement (e.g., the last 12 months) for any duration after intervention commencement,
- *Currently working* — an individual was employed, in any capacity, at time of measurement after intervention commencement,
- *Employment probability* — the probability an individual was employed, at any point, for any duration, after intervention commencement.

Education completion

For education completion some studies reported an outcome capturing 'secondary school, high school or equivalent completion', while others reported 'high school completion' and 'attainment of high school equivalent

qualification' separately. Where these outcomes were reported separately, they were combined — instances where this occurred are detailed in Table 9 in Appendix A.

Consideration of additional active components

During the process of coding components of interventions and their comparators — using the definitions included in Table 1 — it became evident that programmes might also include additional components, beyond those of interest to this review, that may affect outcomes of interest. An additional component 'other' was created to account for their residual contribution.

As 'other' forms a residual category, there is some heterogeneity within it. To explore this, it was further disaggregated into sub-components. To assist in the selection of components focus on, the review team were guided by priority areas nominated by YFF. The selected subcategories included:

- *Case Management* — co-ordination and assistance for participants to access required supports,
- *Paid Work Experience* — temporary experience in a job that is paid, could include job shadowing,
- *Counselling* — includes both job, education and general counselling,
- *Program Access* — including elements that support individuals to participate in the programme i.e., transportation to programme, or provision of childcare,
- *Referral/Brokerage* — including referrals to other services and/or payment for these, or
- *Other* — other components not classified above.

Unit of analysis issues

The unit of analysis for included studies was at the individual level. No unit of analysis issues were identified in the included studies.

Dealing with missing data

For those studies that did not report sufficient data to calculate or transform effect sizes, the study's primary authors were contacted to request the necessary information. Authors of twenty-two (n=22) included studies were contacted to request additional information, five (n=5) of whom responded.

When information was either unavailable or insufficient to calculate an effect size, attempts were made to derive this information based using reasonable assumptions — instances where assumptions were made to interpolate missing data are detailed in Table 9 in Appendix A.

Repeated measures of participants

For studies that reported the same outcome at multiple time points, the review team selected the outcome reported at the last time point, providing us with the longest possible follow-up available across studies.

Quantitative synthesis

The review team quantitatively synthesised included studies using a network meta-analysis (NMA) methodology — see Box 1 for an overview of the method and its key assumptions.⁸ NMA is a statistical technique that can be used to quantitatively synthesise the results from multiple interventions by combining direct and indirect evidence in a network (Tsokani et al., 2022).⁹ Since most of the studies included in this review evaluated programmes that consisted of multiple components (i.e., they were multi-component interventions), we have used an extension to this method — component NMA (CNMA) — that allows us to disentangle the effect of each component.

In its most simple form, NMA is a weighted regression that synthesises both direct evidence (sourced from head-to-head experiments) and indirect evidence (obtained from comparisons that utilise a common comparator) to allow for the comparison of multiple interventions (Petropoulou et al., 2021). There are three major types of NMAs that can be used to disentangle these effects: standard NMA, Additive Component NMA (CNMA) and Interaction Component NMA:

- *Standard NMA* — also known as ‘full-interaction’ NMA. In this analysis, each combination of components identified by the review is considered to be a separate intervention and is assigned its own effect size.
- *Additive Component NMA* — this method assumes that each intervention component has a separate independent effect. Therefore, the total effect of an intervention is equal to the sum of the component effects (this is called the additivity assumption).
- *Interaction Component NMA* — this method extends that additive component NMA by allowing for the inclusion of interactions between two or more pairs (or trios etc.) of intervention components. This means that the total effect can be larger or smaller than the sum of its effects.

Since employment and skills programmes often consist of combinations of these components, a component-NMA method was identified as the preferred method for this review due to its ability to disentangle the relative contribution of each component.

⁸ Quantitative analysis was undertaken using the *netmeta* package for R (Balduzzi et al., 2023).

⁹ Note that studies included in a network meta-analysis need to address the same research question.

Box 1 Overview of Network Meta Analysis and its key assumptions

Network Meta-Analysis

Network Meta-Analysis is a statistical method that allows researchers to compare and rank multiple interventions. Traditional pairwise meta-analyses focused on comparing two interventions at a time, but NMA expands this by incorporating a network of studies that assess different interventions for the same condition or outcome (e.g., A vs. B and B vs. C). By synthesising data from multiple sources, NMA can provide a more comprehensive perspective on treatment effectiveness relative to pairwise meta-analysis.

Component Network Meta-Analysis

Component Network Meta-Analysis takes the concept of NMA a step further by examining the individual components within each intervention. Imagine an intervention as a puzzle, and each component as a unique piece. CNMA allows researchers to study and compare these individual pieces separately, unravelling their specific contributions to the overall treatment effect.

Key Assumptions

There are several important assumptions that underpin both NMA and CNMAs, they include:

Consistency Assumption:

Consistency assumes that the relative treatment effects remain consistent across different comparisons. Essentially, it means that the effectiveness of a particular component within an intervention remains the same regardless of the other components it is combined with or compared against. If this assumption is violated, it suggests that there are factors affecting the results that need to be explored further.

Additivity Assumption:

The additivity assumption posits that the effects of different components within an intervention can be combined in an additive manner. In other words, the overall effect of an intervention is calculated by summing the effects of its individual components. This assumption allows researchers to compare interventions based on the combination of their components, even if those specific combinations have not been directly studied.

Transitivity Assumption:

Transitivity assumes that all treatment comparisons within the network can be connected through a chain of direct or indirect evidence. In other words, it allows for the indirect comparison of interventions. For a NMA to be valid, the assumption of transitivity is crucial. If violated, it suggests that there are systematic differences in the characteristics of the studies or populations, making indirect comparisons unreliable.

Model specification

The review team developed and tested four separate NMA specifications that include different levels of detail about combinations of intervention components and comparators. The specifications were:

- *Specification #1:* Intervention components + consolidated other versus all SAU — in this NMA it is assumed that all studies are compared to services as usual (SAU).
- *Specification #2:* Intervention components + other heterogeneity versus all SAU — this specification is similar to specification #1, in that all studies are assumed to use similar services as usual (SAU), the key difference is that we also factor in the heterogeneity amongst the 'other' category by including these as separate components.

- *Specification #3:* Intervention components + consolidated other versus SAU components + consolidated other — this specification assumes that there may be some heterogeneity amongst services as usual and thus codes SAU using the same components as the intervention, as well as including a consolidated 'other' component.
- *Specification #4:* Intervention components + other heterogeneity versus SAU components + other heterogeneity — the final specification splits both treatment and comparison into "components of interest" and "other components".

Selection of interactions

The review team built and tested interaction-CNMA's for each pair of components present in the NMA treatment composition matrix.

Selection of outcomes for network meta-analysis

The review team considered the feasibility of synthesising each outcome by considering:

- If the reported outcomes assessed the same construct,
- Whether we could transform the reported effect size into a common measure,
- If there were enough studies to populate an NMA (we set a minimum number of ten studies).

If these conditions were met, the process for selecting outcomes for inclusion in the NMA was:

- Reported results were grouped by primary and secondary outcome,
- Where multiple results from the same study were available, the result with the longest follow up time was identified,
- Studies reporting different results for subgroups were identified and their results were pooled to get a population-level result,
- Where required, reported results were transformed to a common effect size using an effect size calculator.

Contextualising results

To support knowledge translation the review team has used two approaches to contextualise the results of our meta-analysis.

Firstly, we have sought to provide some sense of scale to the reported effect sizes. A common approach to interpreting the magnitude of an effect sizes is to apply a set of thresholds proposed by Cohen (1969): 0.2 = small, 0.5 = medium, 0.8 = large. However, Cohen's thresholds were developed from the results of lab-based psychological research and may not accurately characterise the magnitude or importance of an expected effect in a real-

world policy evaluation. At the suggestion of YFF, we have categorised effect sizes by their size into ‘high impact’, ‘medium impact’ and ‘low impact’ categories — see Table 5 for definitions. These ranges were developed by the YFF based on the findings of a recent meta-analysis of active labour market programmes for youth by the International Labour Organization and World Bank (Puerto et al., 2022).

Secondly, to assist readers, we have contextualised the scale of statistically significant, positive effect sizes by translating them to Number Needed to Treat (NNT). NNT is a measure that quantifies the number of individuals who need to receive a specific intervention in order for one additional person to experience the desired outcome compared to an alternative. To derive the NNT, we need to know the rate at which the outcome occurs in the control group — this is called the control group event rate (CER). We calculated this by taking a weighted average of the CER for each study that reported it for employment status (35 studies) and education completion (21 studies). Given that the results varied by outcome we applied different CERs for employment status (0.45) and education completion (0.30).

Table 5 Effect size magnitude categories used in this report

CATEGORY	EFFECT SIZE RANGE (SMD)
High impact	0.2 or greater
Medium impact	Greater than 0.1, but less than 0.2
Low impact	0 to 0.1

Subgroup analysis

The review team undertook multiple subgroup analyses — where sufficient studies were available — to explore how results varied based upon:

- *The location of included studies* — those studies that were conducted in the United States were compared to those conducted in other countries.
- *Study confidence* — those studies in which the review team had high confidence were compared to those with low or medium confidence.
- *Population facing additional barriers* — those that the review team considered to be serving populations facing additional barriers, multiple barriers to employment, or complex needs were compared to those serving populations not known to be facing additional barriers.

Defining additional barriers

Employment and skills programmes are often designed for young people who are identified as being at risk of educational, social and economic disadvantage. However, within that cohort there are some programmes that

are specifically designed for — or serve — particularly vulnerable populations, often due to structural disadvantages and/or additional barriers to employment such as known trauma or disability. We sought to identify these by coding studies based on whether they reported (or were designed to serve):

- *A population living with a disability* — defined as greater than 50 per cent of the population receiving the programme or intervention reporting that they have either a self-identified or diagnosed physical or intellectual disability. Or if the programme was specifically targeted at populations living with a disability.
- *A population with known elevated risks* — defined as greater than 50 per cent of the population receiving the programme or intervention has one or more of the following reported characteristics: current or former experience with the out-of-home care system, self-identified or diagnosed mental health condition, current or former experience with the juvenile justice system, identifies as member of First Nations community, identifies as LGBTIQ+, is a single parent, or if the programme was specifically targeted at serving one of the aforementioned populations.

The review team opted to combine these two groups into a new construct 'young people facing additional barriers'.

Sensitivity analysis

The review team undertook sensitivity analysis that considered the study design. It involved dividing included studies by study design into those that used randomised designs and those that used non-randomised designs. This involved undertaking separate CNMA's for both groups and assessing the variation in results.

Assessing publication bias

Publication bias can arise because studies that have novel or statistically significant findings are more likely to be published. We assessed publication bias by producing a comparison-adjusted funnel plot that plots effect size estimates against a measure of study precision (i.e., standard error), while adjusting for the different number of comparisons present in the network. We examined the symmetry of the plot — deviations from the expected funnel shape can indicate potential publication bias — and performed Egger's test to quantitatively assess the presence of publication bias.

Assessing network coherence

A fundamental assumption of an NMA is that the studies included in the analysis are similar, on average, across important factors that may influence their relative effect — this is referred to as the transitivity assumption. The quantitative version of transitivity is called coherence. The presence of important clinical and/or methodological variation across included studies may be reflected in disagreement between direct and indirect sources of

evidence (Chaimani et al., 2022). When this occurs, it violates the coherence assumption in a network of interventions.

Local incoherence

The Cochrane manual recommends 'Separating Indirect and Direct Evidence' (SIDE) to evaluate 'local incoherence' i.e., incoherence between different combinations of components. A method used to assess SIDE is sometimes referred to in the literature as 'node splitting'. Node-splitting involves separating out the evidence for a particular combination of components into its direct and indirect forms and assessing the discrepancy between them for each combination. This allows for the assessment of the contribution of direct and indirect evidence into each estimate, as well as the ability to test for inconsistency between direct and indirect evidence.

Global incoherence

Incoherence can also occur at the global — i.e., at the network — level. Since a random-effects model was used, we measured global incoherence by fitting a design-by-treatment interaction random effects model (Higgins et al., 2012; Krahn et al., 2013). Incoherence can be assessed through Cochran's Q — aka. Q-test.

Deviations from the protocol

To assist in the reporting of results, we made the following changes in the way some of the outcomes were reported (without changing their meaning):

- “High school or equivalent completion” was used instead of “education completion and qualification (i.e., attainment of secondary-school equivalent education qualification)”, and
- “Vocational education commencement” or “University commencement” was used instead of “access to / in education (i.e., enrolment in TVET or university, or completion of intermediate steps e.g., first year of qualification, progression in TVET)”.

In order to minimise some potential bias that we identified in the process of conducting this review, we undertook some additional analyses not specified in the protocol including:

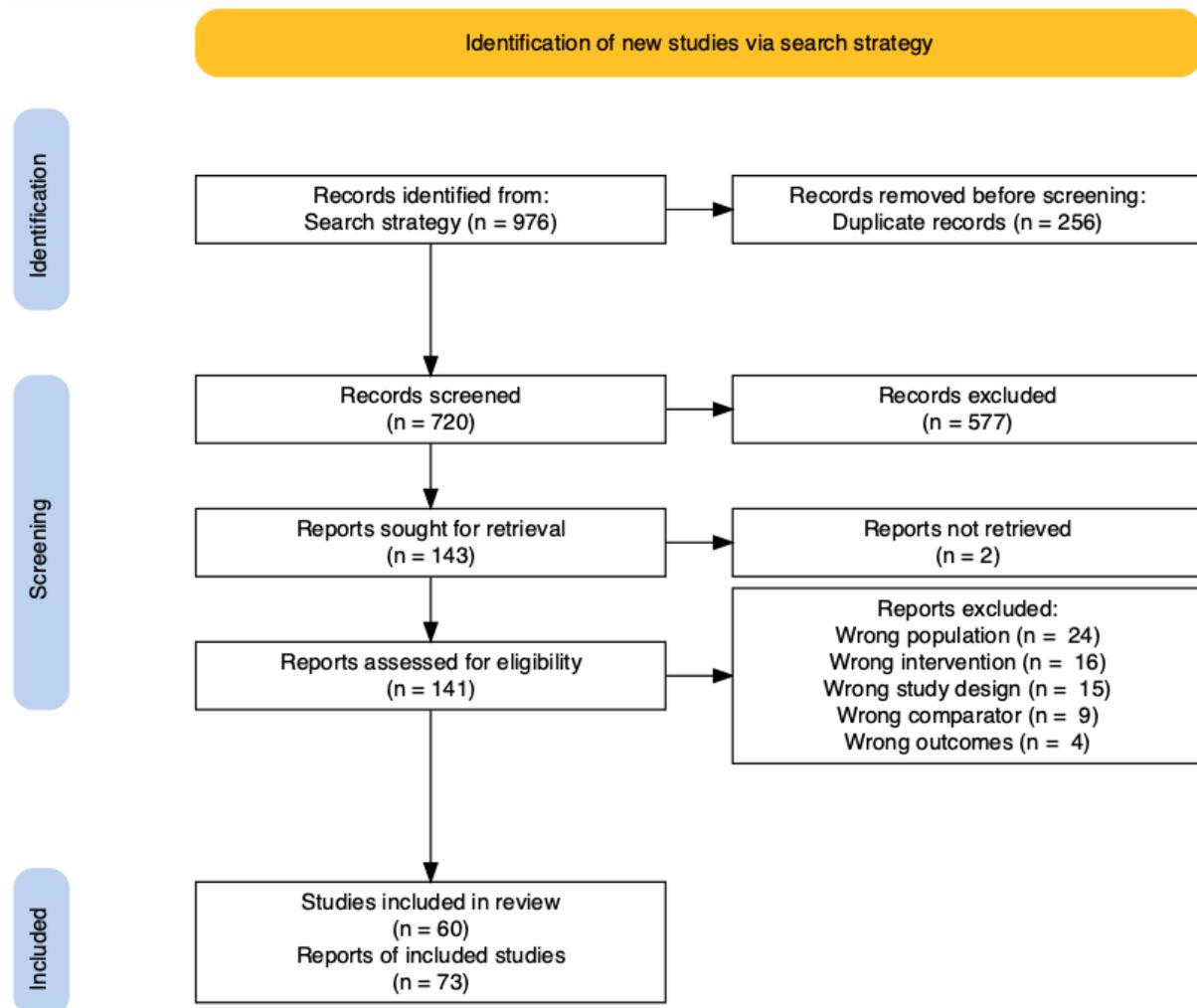
- Subgroup analysis that considered: study confidence (high confidence versus low and medium confidence), location of included studies (United States vs. other locations) and population needs (study population with reported additional barriers versus study population without additional barriers).
- Sensitivity analysis that considered study design (randomised versus non-randomised studies).

Results

Search results

The search strategy yielded a total of 976 records, of which 720 were unique and screened for inclusion. After reviewing titles and abstracts, 141 full-text studies were assessed for eligibility and 60 were included. This process is summarised in the PRISMA flowchart in Figure 1 below.

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart



Characteristics of included studies

Sixty (n=60) studies — reported in 73 articles — met our inclusion criteria and were included in this review. Of these,

- Thirty-five (n=35) papers were different reports of other included studies, studies reported in multiple papers were treated as a single study for the purposes of this review. For studies reported in multiple papers, a primary study was selected to serve as the primary reference – see Table 10 in Appendix A for details.

- Eight (n=8) papers reported analyses of one or more different programmes, which were treated as separate studies (Bloom et al., 1993; Caliendo et al., 2011; Davis & Heller, 2017; T. M. Fraker, Cobb, et al., 2018; Hollenbeck & Huang, 2006, 2016; Maibom et al., 2014; Nadon, 2020)
- Three (n=3) programmes were evaluated at two time points, 10 years apart, involving different populations and were treated as separate studies (Hollenbeck & Huang, 2006, 2016).

Key characteristics of included studies are included in Table 6. A more detailed breakdown of the characteristics of included studies is included in Table 8 in Appendix A. Highlights include:

- Few studies examined apprenticeships (n=2) or on-the-job training (n=8),
- There was a relatively equal distribution of randomised (n=32) and non-randomised (n=28) studies,
- Over two-thirds of included studies were from 'grey' sources (n=41), and
- The rate of publication of studies has increased with half of the included studies published from 2015 onward.

Table 6 Summarised characteristics of included studies

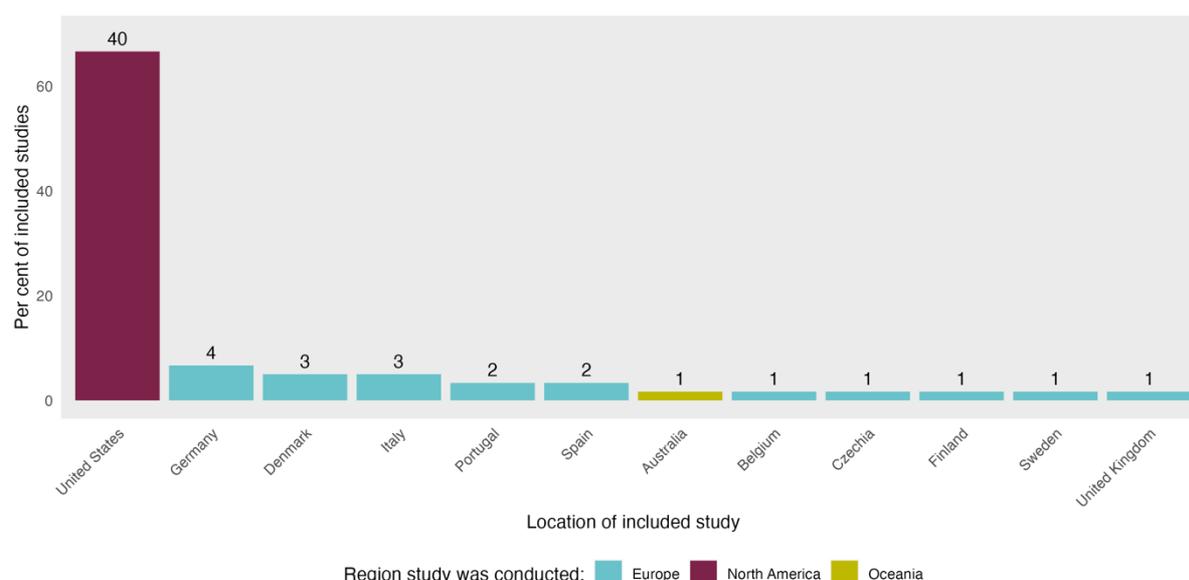
CHARACTERISTIC	NO. OF STUDIES	PER CENT
Intervention or comparison components		
Apprenticeships	2	3.3%
Basic Skills	22	36.7%
Coaching and Mentoring	25	41.7%
Life Skills	20	33.3%
Off-the-job training	22	36.7%
On-the-job training	8	13.3%
Other	53	88.3%
Outcome construct		
Employment status	55	91.6%
Hours worked	23	38.3%
Wages or earnings	38	63.3%
High school or equivalent completion	23	38.3%
Vocational education commencement	5	8.3%

CHARACTERISTIC	NO. OF STUDIES	PER CENT
University commencement	3	5%
Study Design		
Randomised study	32	55%
Non-Randomised study	28	45%
Publication type		
Peer-reviewed publication	19	31.7%
Grey literature	41	68.3%
Year of publication		
1990-1994	4	6.7%
1995-1999	2	3.3%
2000-2004	4	6.7%
2005-2009	6	10%
2010-2014	14	23.3%
2015-2019	26	43.3%
2020-	4	6.7%
Population characteristics		
Facing additional barriers	19	31.7%
Not facing additional barriers	41	68.3%

Location of included studies

Of the sixty included studies, two-thirds (67 per cent, n=40) were conducted in the United States. Of the remaining third (n=20), all but one — which was conducted in Australia — were from Europe. Germany (n=4), Denmark (n=3) and Italy (n=3) were the only countries with more than two included studies — see Figure 2.

Figure 2 Location of included studies



We examined the distribution of components provided as part of interventions delivered in the United States compared to other countries. The results — shown in Table 7 — show both the count of components from included studies and the proportion of evidence that is sourced from the United States. The proportion ranges from 28.6 per cent for off-the-job training through to 100 per cent for apprenticeships, with more than half of the evidence for all components, bar one, sourced from the United States.

Table 7 Distribution of components among included studies by clustered location

COMPONENT	LOCATION		PROPORTION OF EVIDENCE FROM UNITED STATES
	# OF COMPONENTS FROM STUDIES CONDUCTED IN UNITED STATES	# OF COMPONENTS FROM STUDIES CONDUCTED IN OTHER COUNTRIES	
Apprenticeships	2	0	100%
Basic Skills	13	9	59.1%
Coaching and mentoring	20	5	80%
Life Skills	18	2	90%
Off-the-job training	12	8	60%
On-the-job training	2	5	28.6%
Other	35	11	76.1%

Excluded studies

Sixty-eight (n=68) studies were excluded during full-text screening. A selection of these studies and the reasons for their exclusion are detailed in Table 11 in Appendix C.

Included studies not included in NMA

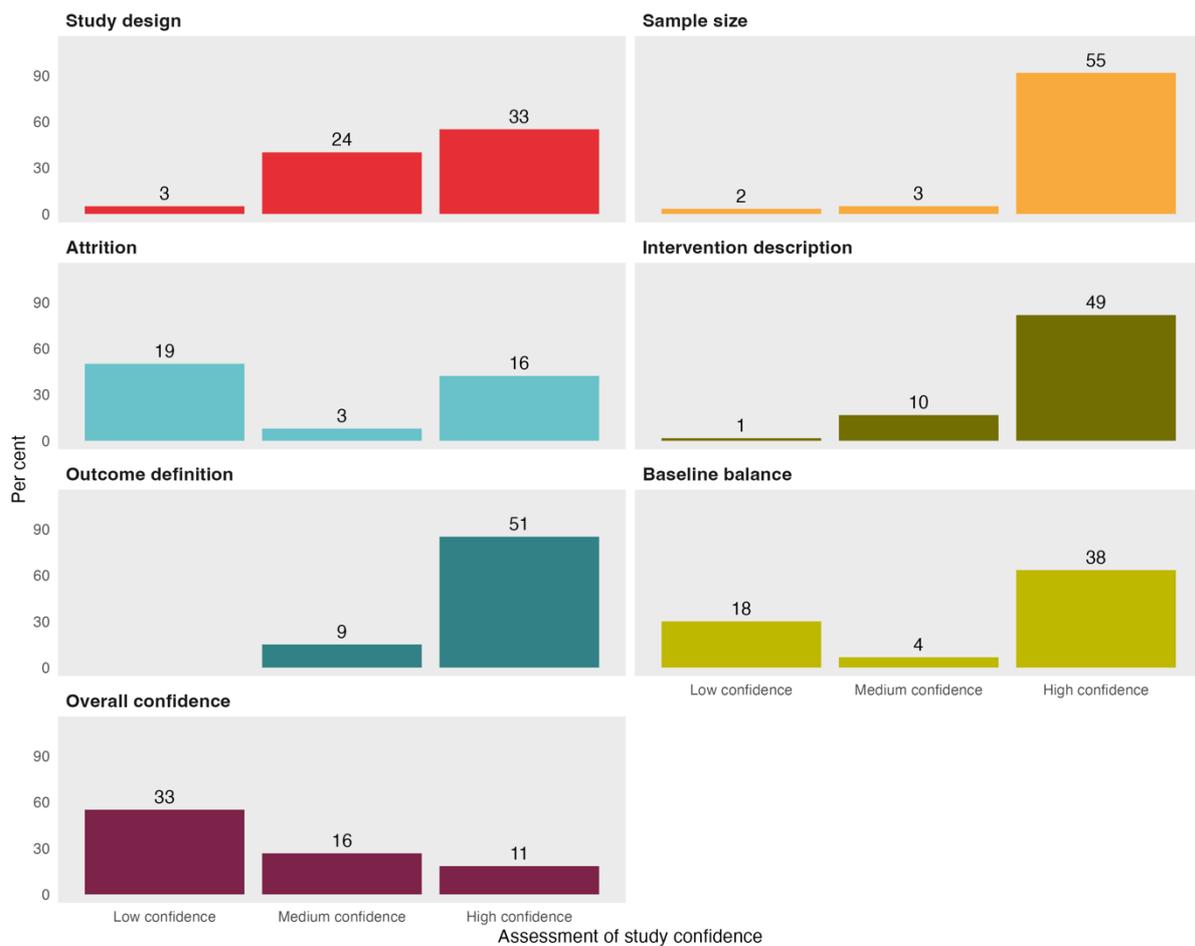
Of the sixty (n=60) included studies, five (n=5) were not included in a network meta-analysis for either employment status or education completion. They were:

- Two studies reported in the same paper by Nadon (2020) were excluded as they used the same dataset as Kim (2019),
- A study by Stromback (2010) only reported wage outcomes,
- A study by Fein & Hamadyk (2018) only reported wage outcomes, and
- A study by Jastrzab (1996) did not report sufficient information to allow for the transformation of results to a common effect size.

Confidence in included studies

Confidence in included studies was assessed using the *Quality assessment of Impact Evaluations* tool (White et al., 2022). The results of the quality assessment are summarised in Figure 3 below. Overall, a majority of studies were assessed as 'low confidence' (55.7 per cent), with the remaining classified as 'medium confidence' (26.2 per cent) or 'high confidence' (18 per cent). Among those studies that we considered to have low confidence, the domains that drove this result were attrition (n=19) and baseline balance (n=18). It is worth noting that this quality assessment tool penalises studies that do not report attrition rates, which — by nature of their design — are not always reported in retrospective, non-randomised studies. Accordingly, the high proportion of studies considered to be 'low confidence' in the attrition domain is partly driven by reporting omissions, which may not necessarily be indicative of concerns surrounding attrition.

Figure 3 Confidence in included studies



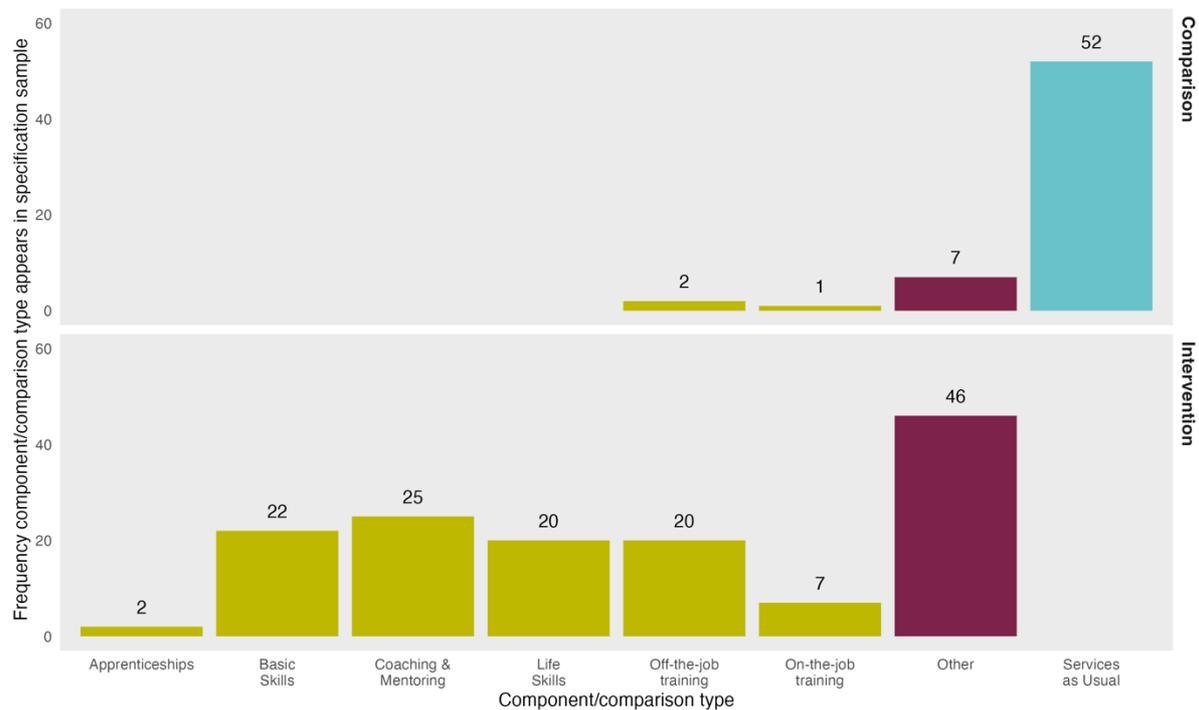
Model selection

The review team developed and tested four separate NMA specifications for all included outcomes — as described in the methodology section — that include different levels of detail about combinations of intervention components and comparators.

To simplify reporting, and support knowledge translation, the results of a single specification (Specification #3) have been reported. In specification #3, treatment interventions are coded as containing one or more of the six components of interest to this review; if the intervention includes an additional active component, it is coded as 'other'. Comparison interventions are coded in the same manner, if no or insufficient information is provided to disaggregate the elements of the comparator it is coded as 'services as usual'.

This was also the specification where the distribution of components allowed for a 'network' to form (for at least one outcome). The distribution of components across all included studies in this network is visualised in Figure 4 below.

Figure 4 Breakdown of components in model three specification



Employment status outcomes

The REA identified fifty-five (n=55) studies that reported sufficient results information to a) derive an individual's employment status and b) transform the result into a common effect size. These results were reported in a range of different outcomes that were grouped into four categories:

- Ever worked following intervention commencement,
- Employed at particular time point following intervention commencement,
- Time to employment following intervention commencement, and
- Employment probability following intervention commencement.

Other employment-related outcomes

The review team also sought to undertake a quantitative synthesis of studies that reported two additional employment-related outcomes:

- Hours worked; and
- Earnings and wages.

Twenty-three (n=23) studies reported a range of measures that would allow the review team to assess time spent in employment. However, of those studies it was only possible to transform three (n=3) of the reported results into a common effect size due to the absence of reported standard errors or standard deviations.

Thirty-eight (n=38) studies reported results that allowed the review team to assess how much participants earned. Unfortunately, it was only possible to

transform twelve (n=12) of these results into a common effect size. As with the hours worked outcome, the issue preventing the review team from transforming these results was the absence of reported standard errors or standard deviations.

As a result, it was not possible to undertake an NMA for either of these two outcomes. Details of included studies for hours worked are included in Table 12 and for wages and earnings in Table 13 in Appendix C.

Structure of network for employment status

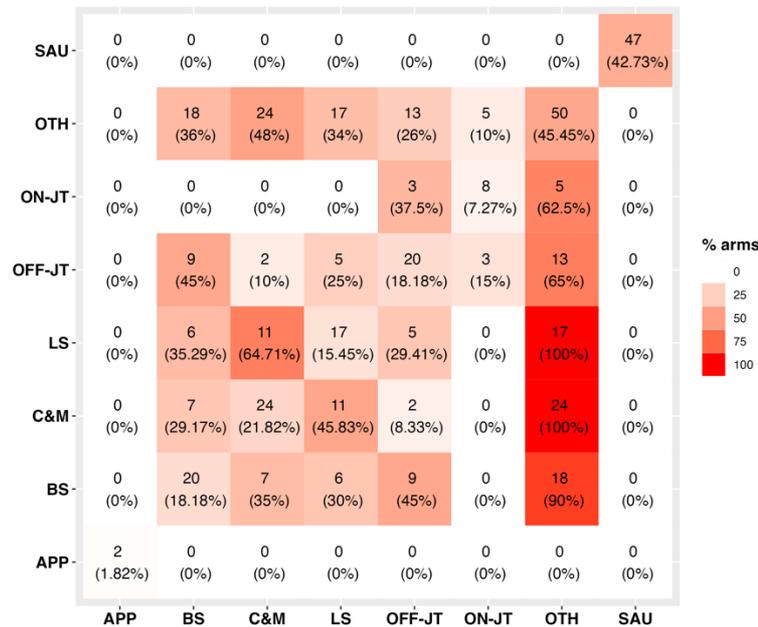
The relationships between different components, and combinations of components, are visualised in a network map. Each node represents a unique combination of components and the linkages (or edges) between them. Line thickness represents the relative size of the linkage between each node. The network map for the employment status NMA is visualised in Figure 5. The map shows that the network is 'fully connected' (i.e., all nodes are linked), that there are 20 unique treatments (i.e., combinations of components) in the network and 23 designs (i.e., edges between them). The figure also shows the number of studies that are included in each pairwise comparison.

The network map shows that the majority of designs share a common comparator — services as usual. However, there are seven pairwise comparisons¹⁰ that do not, and these help form additional connected loops within the network.

¹⁰ a) Coaching & Mentoring+Other:Other, b) Apprenticeships:Other, c) Coaching & Mentoring+Other:Other, d) Life Skills+Off-the-job training+Other:Off-the-job training, e) Life Skills+Coaching & Mentoring+Other:Other, f) Basic Skills+Life Skills+Coaching & Mentoring+Other:Off-the-job training+On-the-job Training+Other, g) Off-the-job training: Other.

'life skills' was seen in 45.8% of interventions that included 'coaching and mentoring'.

Figure 6 Heatmap showing the distribution of components among included studies for employment status



Total number of study arms: 110

Assessing the impact of combinations of components on employment status

Prior to fitting a component meta-analysis, we fitted a standard network meta-analysis. This type of analysis does not separate the components individually; instead it looks at the effect of different combinations of components as they were delivered i.e., as part of larger programmes.

Fifty-five (n=55) studies — 18 per cent (n=10) of which we had high confidence in — that compared 20 different combinations of components were included in a standard NMA. A random-effects model was selected on the basis that this was the conservative option due to the potential presence of unobserved heterogeneity within both the included studies and study populations (i.e., to account for the fact that included studies may vary in design and method in a manner that has not been measured).

A moderate degree of heterogeneity ($\tau = 0.180$) was detected among the effect sizes of the combinations of components. A high level of inconsistency ($I^2: 96.7\%$, 95% CI: [96.1%; 97.2%]) was also identified suggesting substantial variability in effect estimates beyond what could be attributed to chance and that studies included in the analysis differ significantly in their outcomes. Consequently, caution should be exercised when interpreting the effect estimates, as the observed heterogeneity may impact the generalisability of the results.

The results of the standard NMA — depicted in Figure 7 below — show that there are five combinations of components that, when delivered together, all show a statistically significant and high impact on employment status relative to services as usual. Ordered from largest to smallest they are:

- Two (n=2) studies included in the network combined On-the-job training + Other ($g=0.48$, 95% CI: [0.11, 0.84], $p < 0.01$). This combination has an NNT of 5.3 (95% CI: 3.2-22.9) indicating that, on average, for every five individuals who received the intervention one additional individual will subsequently be employed.
- Five (n=5) studies included in the network combined Basic Skills + Off-the-job training + Other ($g=0.30$, 95% CI: [0.12, 0.48], $p < 0.01$). The NNT for this combination is 8.4 (95% CI: 5.3-20.9), meaning that, on average, for every eight individuals who received the intervention, one additional individual will subsequently be employed.
- Three (n=3) studies included in the network only included On-the-job training ($g=0.25$, 95% CI: [0.05, 0.46], $p < 0.01$). The NNT for this combination is 10.1 (95% CI: 5.5-50.4), indicating that, on average, for every ten individuals who receive the interventions, one additional individual will be employed.
- Six (n=6) studies included in the network combined Life Skills + Coaching & Mentoring + Other ($g=0.24$, 95% CI: [0.08, 0.39], $p < 0.01$). With an NNT of 10.5 (95% CI: 6.5-31.5) means that, on average, for every ten individuals who receive the interventions, one additional individual will be employed with this combination.
- Four (n=4) studies included in the network only included Off-the-job training ($g=0.23$, 95% CI: [0.06, 0.40], $p < 0.01$). The NNT for this combination is 10.9 (95% CI: 6.3-42), meaning that, on average, for every eleven individuals who receive the intervention, one additional individual will be employed.

This analysis should be seen as complementary to the CNMA. While it cannot disentangle the individual effect of a component, it can provide some insight into the impact of combinations of components when they are delivered together.

Non-significant findings of interest for employment status

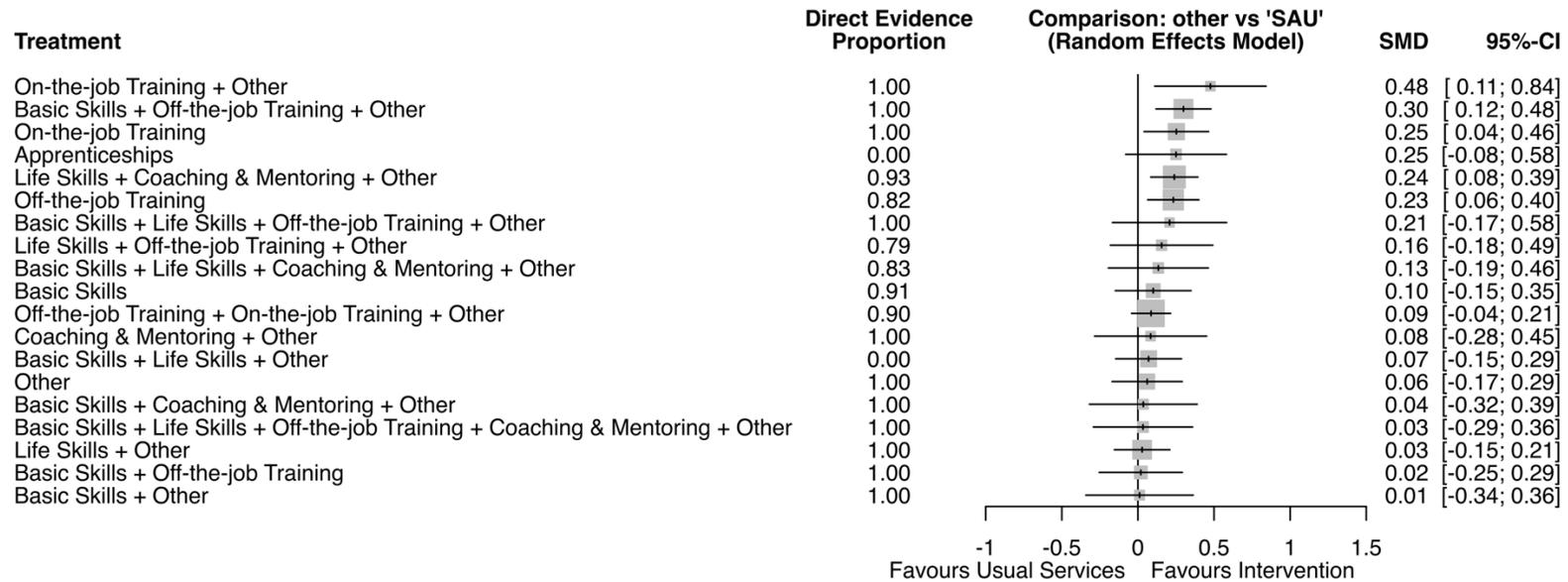
There is a risk that the network was underpowered to detect statistically significant differences in some combinations due to the inclusion of a limited number of studies investigating these combinations. Considering this, we have highlighted some combinations that are not statistically significant as they may translate into significant effects in future analyses as the number of studies with similar findings increase. They include:

- Two (n=2) studies that examined Apprenticeships ($g=0.25$, 95% CI: [-0.08, 0.58], $p > 0.05$).

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

- One (n=1) study that included Basic Skills + Life Skills + Off-the-job training + Other ($g=0.21$, 95% CI: $[-0.17, 0.58]$, $p > 0.05$).

Figure 7 Forest plot depicting results of a standard NMA for employment status



Assessing the impact of each component on employment status

One of the main goals of this review is to assess the relative contribution of each of the program components toward achievement of the outcome. The influence of individual components can be evaluated in an additive CNMA model assuming that the effect of each combination is the sum of the effects of its components (a.k.a. the additivity assumption).

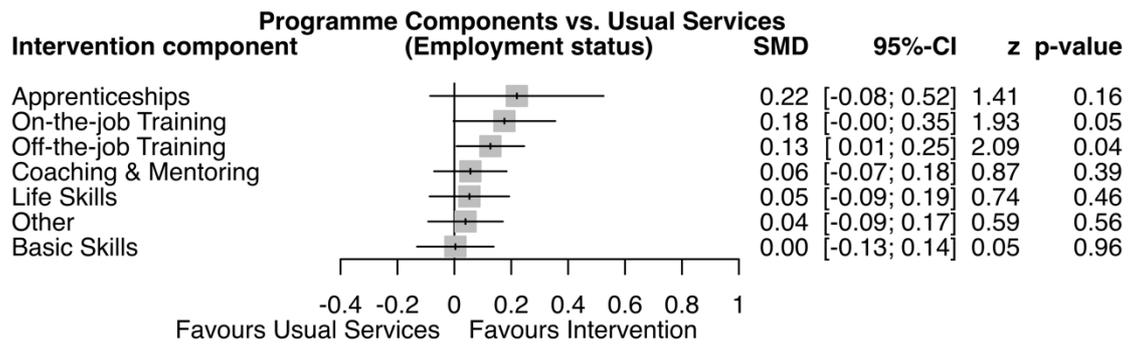
All of the CNMAs use the same inputs as those used in the standard NMA i.e., fifty-five (n=55) studies — 18 per cent (n=10) of which we had high confidence in. However, instead of assessing the impact of combinations of components, the component NMA assesses the incremental effect of an individual component relative to services as usual.

The results — depicted in Figure 8 below — showed that individuals in included studies that received:

- Off-the-job training ($g=0.13$, 95% CI: [0.01; 0.25], $p < 0.05$) had statistically significant, moderate sized impact — meaning those who received off-the-job training were more likely to attain employment than those who received services as usual. To put this in context, this means that for every 19.3 (95% CI: 10.1-252.5) individuals who receive this component, one additional person will be employed.
- The effect of both apprenticeships ($g=0.22$, 95% CI: [-0.08; 0.52], $p > 0.05$) and on-the-job-training ($g=0.18$, 95% CI: [-0.00; 0.35], $p > 0.05$) is not statistically significant, however we report them here because there are some indications that the network may be under-powered for detecting small, but meaningful differences, and findings may translate into significant effects in future analyses as the number of studies with similar findings increase.
- All other components — Coaching and Mentoring, Life Skills, Basic Skills and Other — had small effects that were not statistically significant.

Similar to the standard NMA, a moderate degree of heterogeneity ($\tau = 0.197$) was detected among the assessed components. A high level of inconsistency ($I^2: 97.8\%$, 95% CI: [97.5%; 98.1%]) was also identified, suggesting a substantial variability in effect estimates beyond what could be attributed to chance and that studies included in the analysis differ significantly in their outcomes. Consequently, caution should be exercised when interpreting the effect estimates, as the observed heterogeneity may impact the generalisability of the results.

Figure 8 Forest plot showing the results of an additive CNMA for employment status



Assessing the impact of interactions between components on employment status

There might be instances where delivering multiple components simultaneously may have a multiplicative effect on outcomes for participants. In these situations — where the effect of two components is greater than the sum of their parts — the additivity assumption is not met. It is possible to test whether or not components have a multiplicative effect by adding an interaction term into the NMA model. In the event that no significant interactions are identified we can conclude that the additivity assumption is reasonable.

To test this, we sought to identify if any combinations of components had a multiplicative effect by fitting a series of interaction models. In deciding what combinations of components to interact we were constrained by necessity of each component to be present within the interventions included in this analysis.

For employment status, we fitted interaction models for every two-way combination of components present in the network. They were:

- Basic Skills x Off-the-job training
- On-the-job training x Other
- Basic Skills x Other
- Life Skills x Other
- Coaching & Mentoring x Other

None of the tested interactions were significant at the $p < 0.05$ level. Based on this, there is no evidence of interactions between them. However, it is unclear whether this is due to a lack of statistical power or the absence of an interactive effect. Forest plots showing the results of these analyses are presented below.

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

Figure 9 Forest plot depicting results of CNMA with interaction between Basic Skills x Off-the-job training for employment status

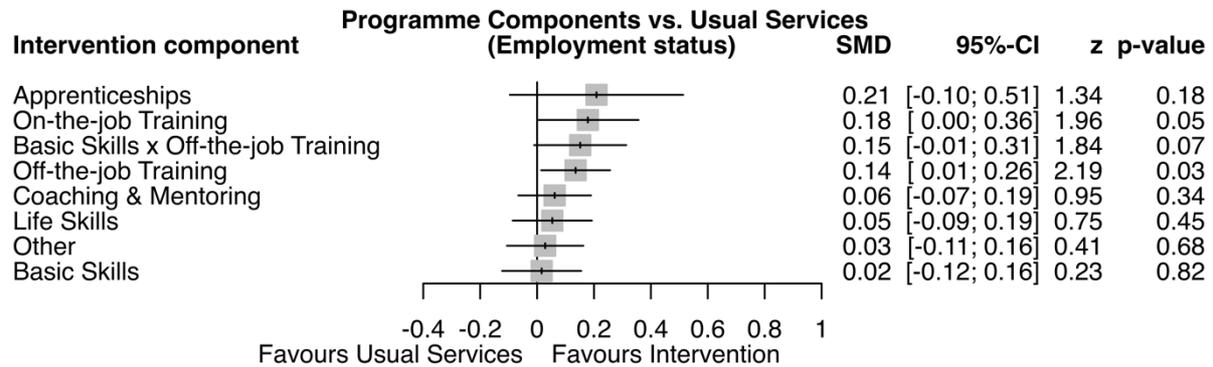


Figure 10 Forest plot depicting results of CNMA with interaction between On-the-job training x Other for employment status

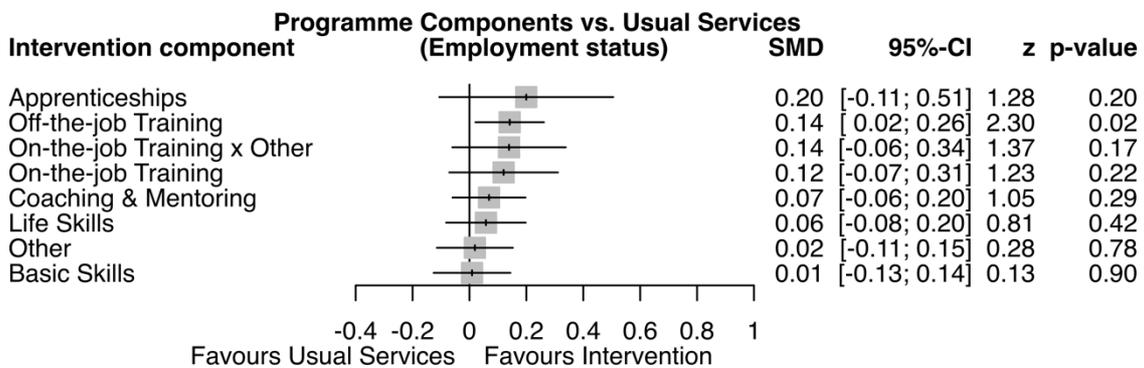


Figure 11 Forest plot depicting results of CNMA with interaction between Basic Skills x Other for employment status

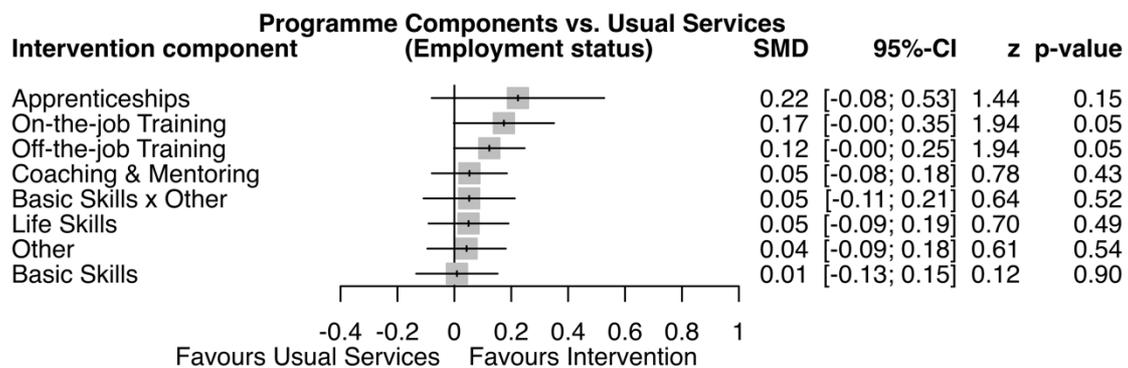


Figure 12 Forest plot depicting results of CNMA with interaction between Life Skills x Other for employment status

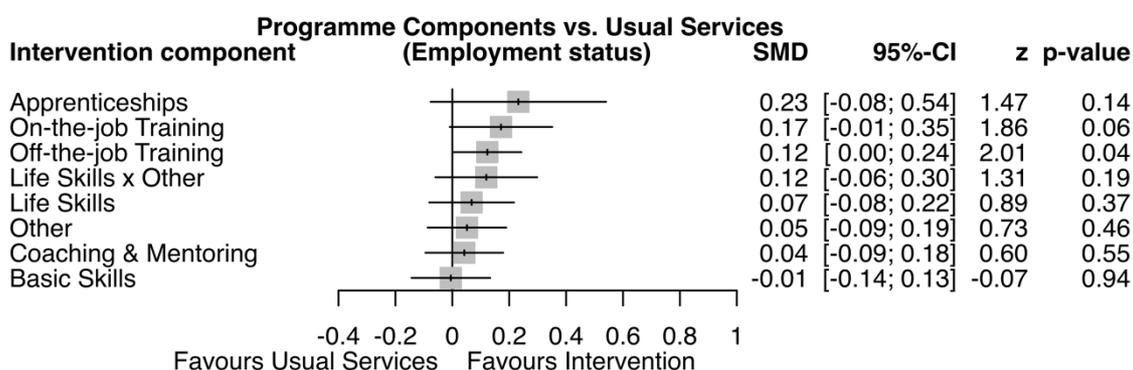
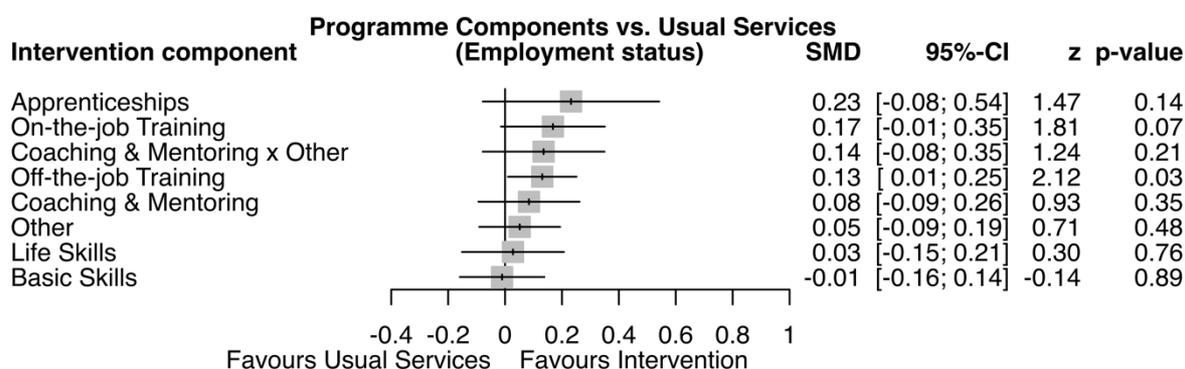


Figure 13 Forest plot depicting results of CNMA with interaction between Coaching & Mentoring x Other for employment status



Identification of a preferred model for employment status

The results of the standard NMA, additive and interaction CNMAs are summarised in Figure 29 in Appendix D. These results suggest that the additive approach provides increased precision, relative to the standard approach. However, no additional precision appears to be provided through the use of an interaction approach. As a result, the review team have identified that the additive CNMA is our preferred specification to establish component effects (compared to the interaction CNMA). However, it is important to note that the standard NMA provides additional supplementary information about promising combinations of components that might warrant exploration.

Education outcomes

The REA identified twenty-three (n=23) studies that reported on outcomes that identified whether an individual completed high school or attained an equivalent qualification (hereafter referred to as 'education completion'). Three types of outcomes were reported:

- Completion of high school,
- Attainment of equivalent qualification, or
- Completion of high school or equivalent qualification.

For our analyses, these three outcomes were combined to create 'high school or equivalent qualification'. A common effect size was calculated for each of the twenty-three studies reporting this outcome.

Quantitative synthesis was not possible for studies reporting two additional education-related outcomes:

- Vocational education commencement, or
- University commencement.

Five (n=5) included studies reported a range of outcome measures that represented whether an individual commenced vocation education. Due to the small number of included studies, transformation of these results into a common effect size was not possible — details of these studies are included in Table 14 in Appendix C.

Three (n=3) studies reported outcomes representing whether a participant commenced university studies. Again, there were not enough studies to undertake an NMA so the transformation of these results into a common effect size was not possible — details of these studies are reported in Table 15.

Structure of network for education completion

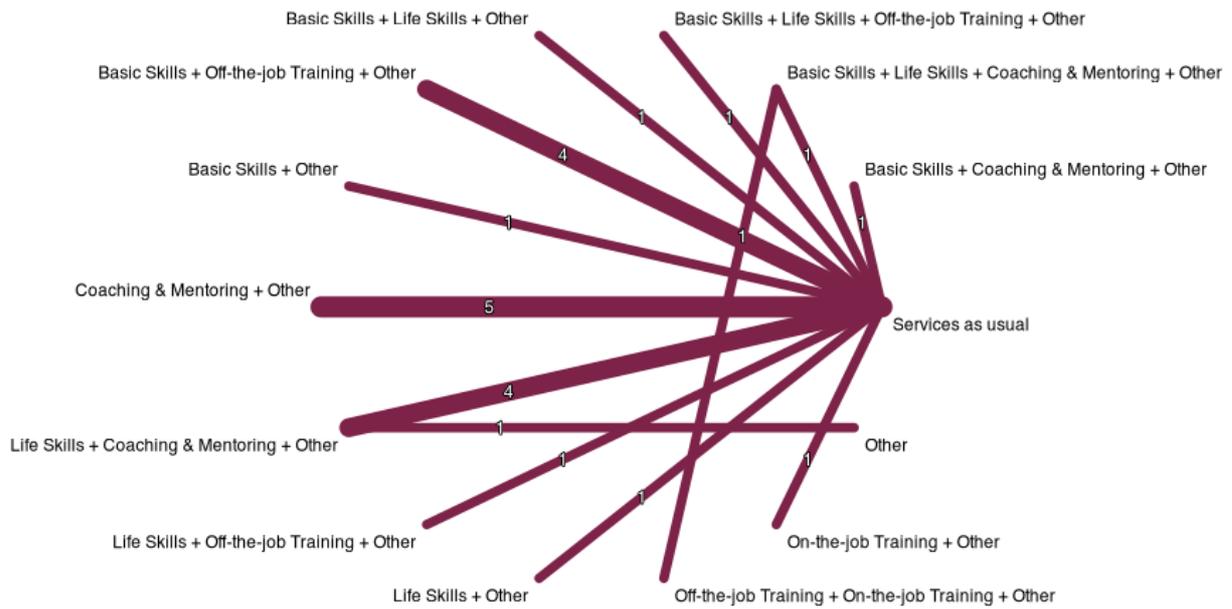
The relationships between different combinations of components that made up the interventions that were evaluated in the included studies reporting education completion outcomes are visualised in Figure 14 below. The map shows that the network is 'fully connected' (i.e., all nodes are linked), that there are 14 unique treatments (i.e., combinations of components) in the network and 13 designs (i.e., edges between them).

Relative to the network map for employment status visualised in Figure 5, an important difference is that there are no connected loops present amongst studies reporting education completion outcomes.

The implication of this is that the evidence derived from this network is either completely direct (e.g., Basic Skills + Other vs. Services as usual) or completely indirect (e.g., Coaching & Mentoring + Other can only be compared with Basic Skills + Other via its common comparator Services as usual). Without the use of mixed evidence (which comes from both direct and indirect sources) it is not possible to determine how consistent the network is either at the local level (between treatments) or globally (across the entire network).

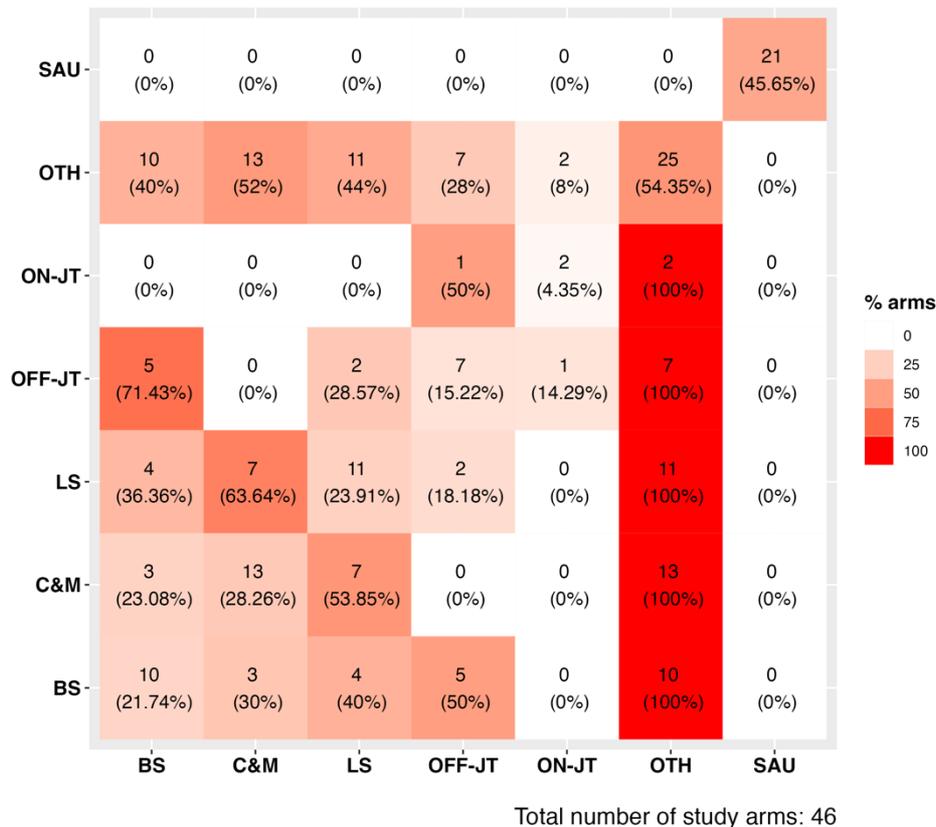
A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

Figure 14 Network map for education completion NMA



The component crosstabulation for employment or skills programme that report education completion outcomes is visualised in Figure 15. The figures reported in the 'diagonal elements' show that the most frequently occurring active components — i.e., excluding 'other' (n=25) and 'services as usual' (n=21) — are 'coaching and mentoring' (28.3 per cent) which was observed in 13/46 intervention arms, followed by 'life skills' (23.9 per cent) seen in 11/46.

Figure 15 Heatmap showing distribution of components among included studies for education completion¹¹



Assessing the impact of combinations of components on education completion

As with the employment status outcome, the review team fitted a standard network meta-analysis that looked at the impact of differing combinations of components on education completion outcomes.

Twenty-three (n=23) studies — 39 per cent (n=9) of which we had high confidence in — that compared 13 different designs (i.e., combinations of components) were included in a standard NMA. A random-effects model was selected on the basis that it was the conservative option used due to the potential presence of unobserved heterogeneity within both the included studies and study populations.

A moderate degree of heterogeneity ($\tau = 0.176$) was detected among the effect sizes of the combinations of components. A high level of inconsistency ($I^2: 81.2\%$, 95% CI: [67.3%; 89.1%]) was also identified suggesting a substantial variability in effect estimates beyond what could be attributed to chance and that studies included in the analysis differ significantly in their outcomes. Consequently, caution should be exercised when interpreting the effect

¹¹ **Plot legend** — BS: Basic Skills, LS: Life Skills, OFF-JT: Off-the-job training, ON-JT: On-the-job-training, APP: Apprenticeships, C&M: Coaching and mentoring, OTH: Other (residual) component.

estimates, as the observed heterogeneity may impact the generalisability of the results.

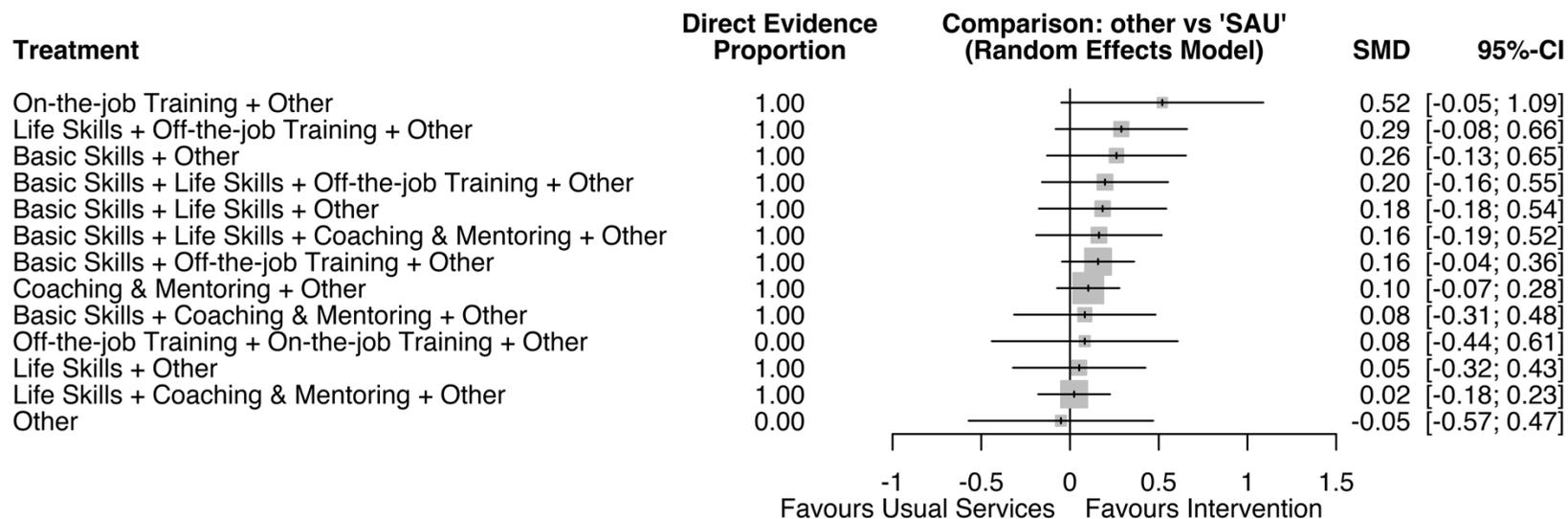
The results of the standard NMA — presented in Figure 16 — indicate the none of the different combinations of components had a statistically significant effect on the completion of high school or equivalent education, which means that we cannot be confident observed differences in outcomes are not due to chance alone (i.e., the evidence is weak).

Non-significant findings of interest for education completion

There is a risk that the network was underpowered to detect statistically significant differences in some combinations due the inclusion of a limited number of studies investigating these combinations. Considering this, we have highlighted some combinations of components that do not report statistically significant results as they may translate into significant effects in future analyses as the number of studies with similar findings increase. They are:

- One (n=1) study included the components On-the-job training + Other ($g=0.52$, 95% CI: [-0.05, 1.09], $p > 0.05$);
- One (n=1) study included Life Skills + Off-the-job training + Other ($g=0.29$, 95% CI: [-0.08, 0.66], $p > 0.05$) components;
- One (n=1) study included Basic Skills + Other components ($g=0.26$, 95% CI: [-0.13, 0.65], $p > 0.05$); and
- One (n=1) study included Basic Skills + Life Skills + Off-the-job training + Other ($g=0.20$, 95% CI: [-0.16, 0.55], $p > 0.05$) components.

Figure 16 Forest plot depicting results of a standard NMA for education completion



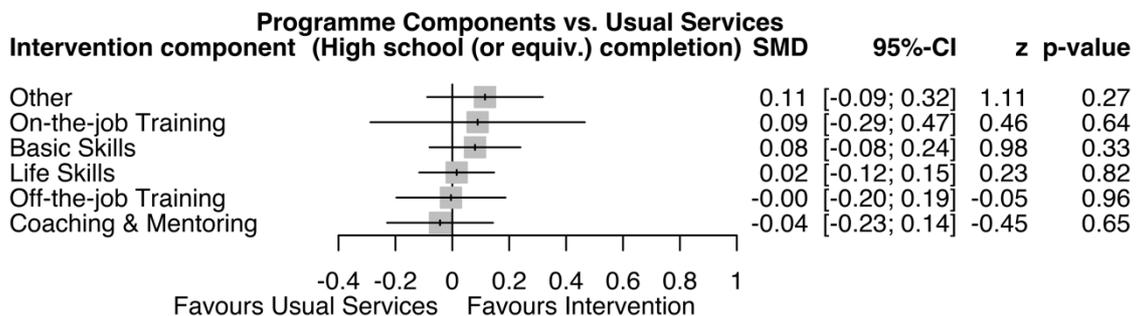
Assessing the impact of each component on education completion

An additive component NMA allows the assessment of the relative contribution of each of the intervention components toward achievement of the outcome, assuming the additivity assumption holds. As it is built on the standard NMA it uses the same inputs i.e., twenty-three (n=23) studies — 39 per cent (n=9) of which we had high confidence in.

Similar to the standard NMA, a moderate degree of heterogeneity ($\tau = 0.130$) was detected among the assessed components. A high level of inconsistency ($I^2: 73.7\%$, 95% CI: [58.1%; 83.5%]) was identified suggesting a substantial variability in effect estimates beyond what could be attributed to chance and that studies included in the analysis differ significantly in their outcomes. Consequently, caution should be exercised when interpreting the effect estimates, as the observed heterogeneity may impact the generalisability of the results.

The results of an additive CNMA for education completion — presented in Figure 17 — suggest that study participants who received any of the individual components were no more likely to complete high school, or receive an equivalent qualification, than those individuals who received services as usual.

Figure 17 Forest plot showing the results of an additive CNMA for education completion



Assessing the impact of interactions between components on education completion

For education completion, the review team identified four potential pairs of components that could be combined in an interaction. They were:

- On-the-job training x Other
- Basic Skills x Other
- Life Skills x Other
- Coaching & Mentoring x Other

As with employment status, there were no significant differences in the impact of the components individually or interacting in any of these models — these results are presented in the forest plots below.

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

Figure 18 Forest plot depicting results of CNMA with interaction between Coaching & Mentoring x Other for education completion

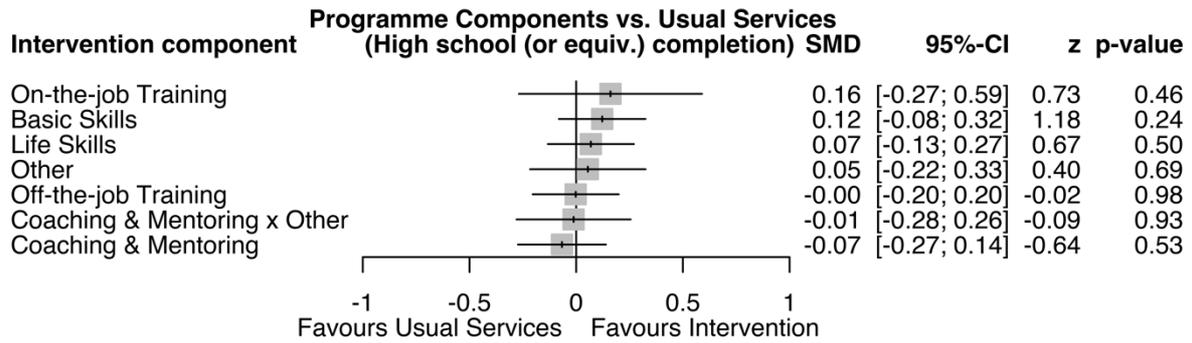


Figure 19 Forest plot depicting results of CNMA with interaction between Life Skills x Other for education completion

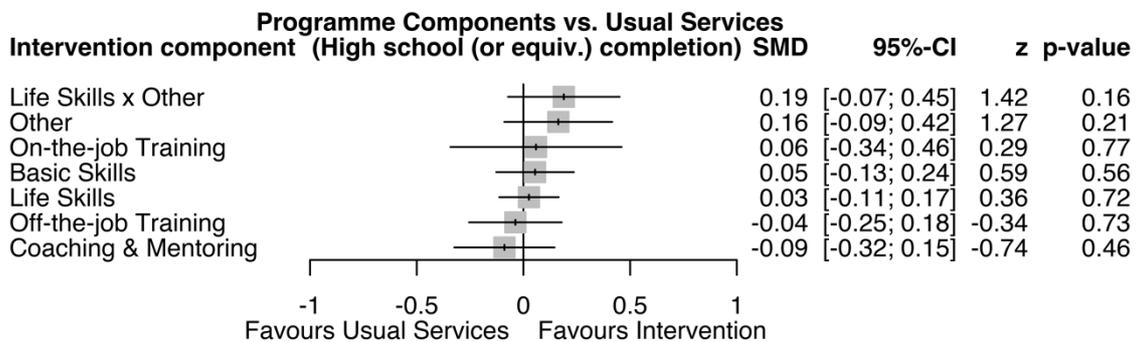


Figure 20 Forest plot depicting results of CNMA with interaction between Basic Skills x Other for education completion

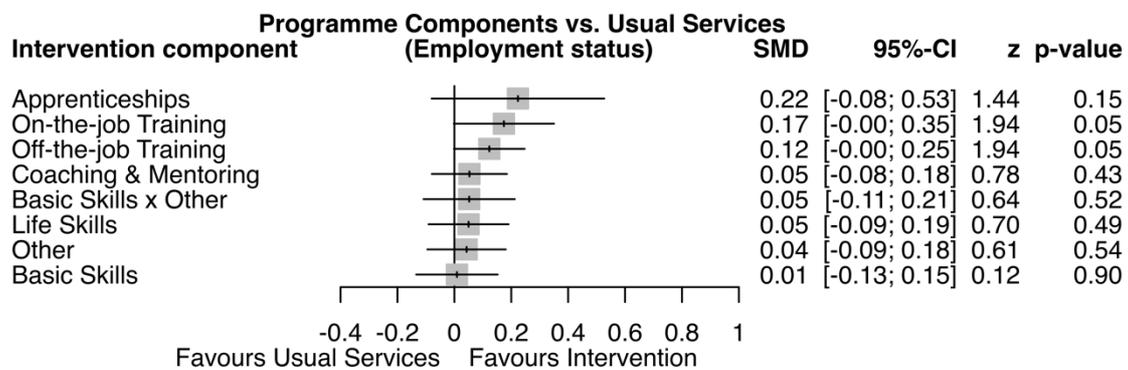
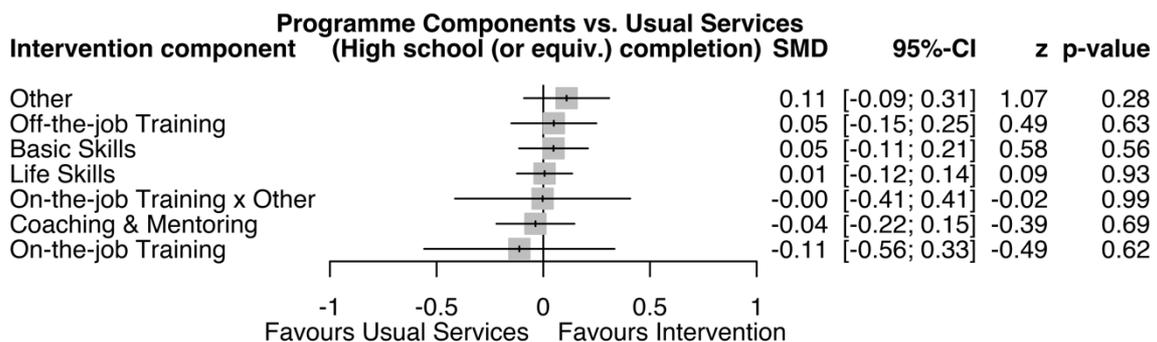


Figure 21 Forest plot depicting results of CNMA with interaction between On-the-job training x Other for education completion



Identification of a preferred model for education completion

The results of the standard NMA, additive and interaction CNMAs for education completion are summarised in Figure 30 in Appendix D. As with the employment status analysis, the results suggest that the additive approach provides increased precision, relative to the standard approach. However, no additional precision appears to be provided through the use of an interaction approach. As a result, the review team have identified that the additive CNMA is our preferred specification to establish component effects (compared to the interaction CNMA). However, it is important to note that the standard NMA provides additional supplementary information about promising combinations of components that might be worth exploring.

Subgroup analysis

A series of subgroup analyses were undertaken in order to test the sensitivity of the results of the additive CNMA for both employment status and education completion outcomes.

Study confidence

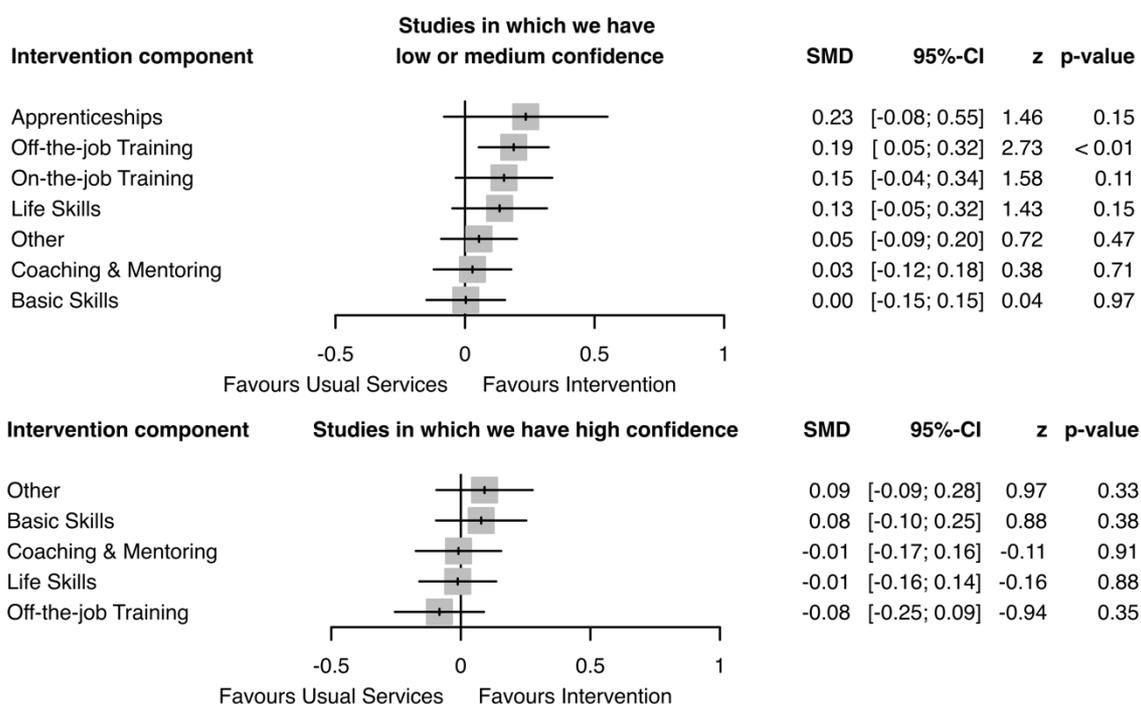
To assess the impact of study confidence, the review team compared the results of an additive CNMA that was limited to studies that were rated as 'high confidence' to the results of an additive CNMA of studies that were rated 'low or medium confidence'.

Subgroup analysis of employment status by study confidence

Among studies that reported employment status outcomes, sixteen (n=16) were assessed as 'high confidence', and forty-five (n=45) were 'low' or 'medium confidence'.

Figure 22 presents this analysis. There were no significant differences between the two groups. While off-the-job training for low to medium confidence studies appears significant, this is most likely to be the result of a reduction in statistical power due to the relatively small number of studies considered 'high confidence'.

Figure 22 Forest plot depicting the result of subgroup analysis for employment status by study confidence



Subgroup analysis of education completion by study confidence

Of the twenty-three (n=23) included studies reporting education completion outcomes, nine (n=9) were assessed as 'high confidence', and fourteen (n=14) were 'low or medium confidence'. In this case, there were insufficient studies to undertake a subgroup analysis.

Study location

As observed in Figure 2, more than two-thirds of included studies were conducted in the United States. There are aspects of both the training and labour market in the United States that may not be generalisable to other countries. For example:

- Relative to other countries in which included studies were conducted, the labour market tends to be more 'flexible' in the United States, which can affect the ease of attaining employment (and also its length of tenure and the relative advantage conferred by additional training on an individual's wage level), and
- General services available to young people not in employment, education or training — which constitute 'services as usual' — may be more limited in the United States relative to those available in other countries.

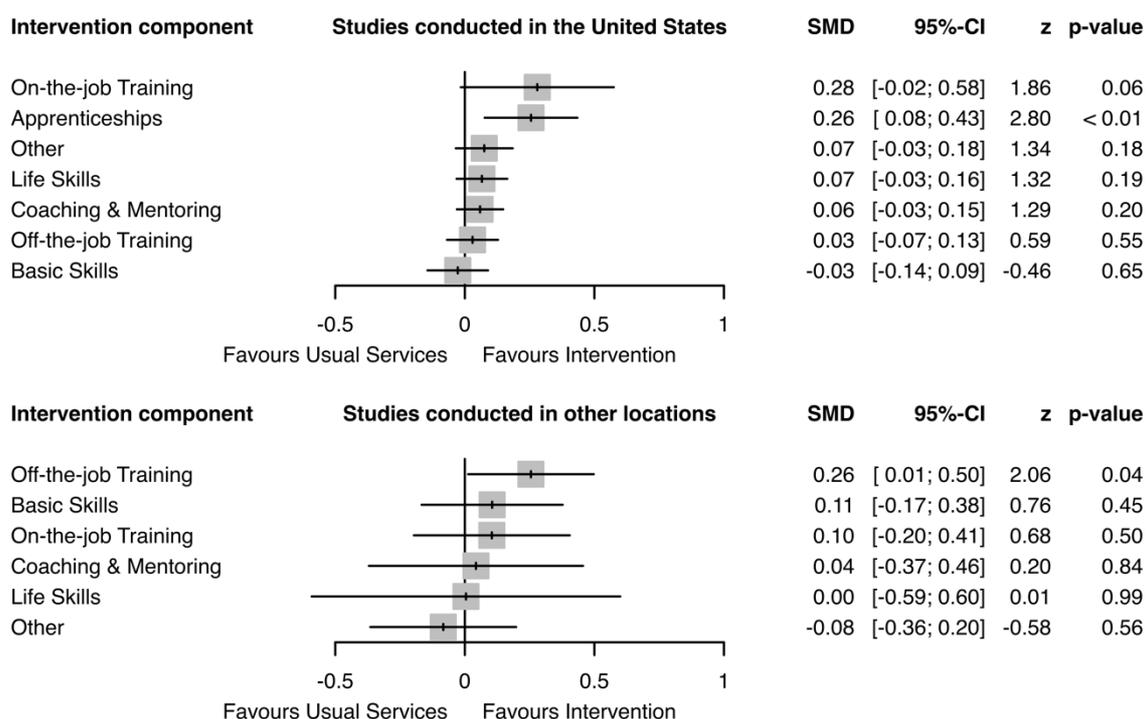
For the reasons outlined above the review team sought to test whether there were any significant differences between studies conducted in the United States, relative to those conducted in other countries.

Subgroup analysis of employment status by study location

Of the fifty-five (n=55) included studies that report employment status outcomes, thirty-six (n=36) were conducted in the United States, with the remaining nineteen (n=19) conducted elsewhere.

The results — presented in Figure 23 below — indicate that apprenticeships have a small but significant impact on employment, however they have only been evaluated in the United States. Therefore, it is unclear whether these findings are generalisable to other countries.

Figure 23 Forest plot depicting the result of subgroup analysis for employment status by study location



Subgroup analysis of education completion by study location

Of the twenty-three (n=23) included studies reporting education completion outcomes, twenty-two (n=22) were from the United States, which did not allow for subgroup analysis by study location.

Study population

While unemployed or out-of-school young people typically experience some element of disadvantage, some subgroups within this population face additional barriers. This could potentially affect the results of this review in a number of ways. Young people presenting with additional barriers may not respond to a programme component in the same way young people without these barriers might — this could lead to results favouring the null hypothesis (i.e., favouring services as usual). Alternatively, the nature of services as usual — i.e., if there is no support available to the comparison

group, or little change is expected in their outcomes without support — could bias results in favour of the alternative hypothesis (i.e., favouring treatment).

In order to test this, the review team undertook separate additive CNMA's that involved young people with reported additional barriers and those involving young people without reported additional barriers.

Subgroup analysis of employment status by study population

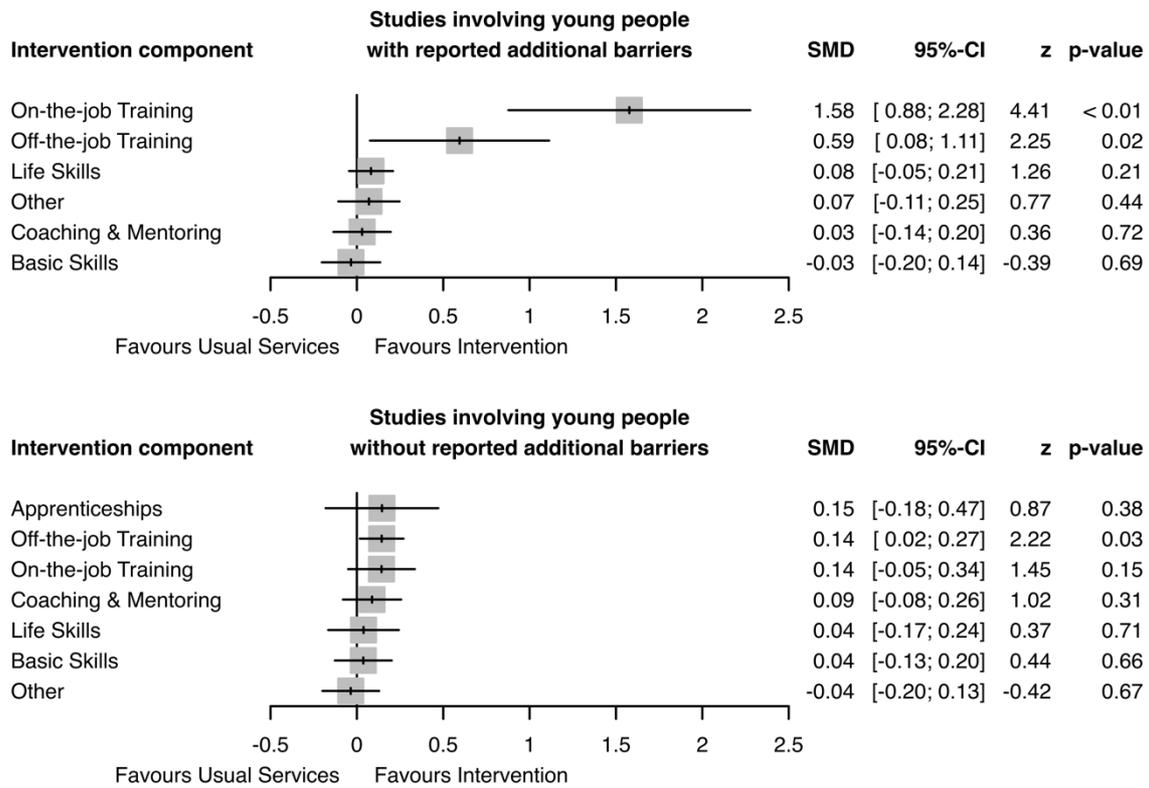
Of the fifty-five (n=55) included studies that report employment status outcomes, seventeen (n=17) involved populations with reported additional barriers, with the remaining thirty-eight (n=38) involving populations that do not report additional barriers.

The results — depicted in Figure 24 — show some marked differences between the two groups. Amongst those who report additional barriers, on-the-job training ($g=1.58$, 95% CI: [0.88-2.28], $p < 0.01$) and off-the-job training ($g=0.59$, 95% CI: [0.08-1.11], $p < 0.05$) both report very high impact effect sizes. Amongst young people who do not report additional barriers, off-the-job training ($g=0.14$, 95% CI: [0.02-0.27], $p < 0.05$) has a statistically significant medium impact on employment status.

To put these results in context, for every 2.1 (95% CI: 1.9-3.1) young people who face additional barriers who receive on-the-job training, on average, one additional young person would be expected to subsequently attain employment. On average, for every 4.4 (95% CI: 2.6-31.5) young people facing additional barriers who receive off-the-job training, it is expected that one additional young person will subsequently achieve employment. For young people who do not report facing additional barriers, on average, for every 17.9 (95% CI: 9.3-126.2) who receive off-the-job training one additional young person will subsequently be employed.

The key point emerging from this analysis is that both on-the-job training and off-the-job training have very large and large effects on employment outcomes for young people who report that they face additional barriers. Another way to look at this, is that young people who do not face additional barriers are more likely to be able to find employment on their own, while those with additional barriers appear to benefit from the assistance of some of these components.

Figure 24 Forest plot depicting the result of subgroup analysis for employment status by study population with reported additional barriers

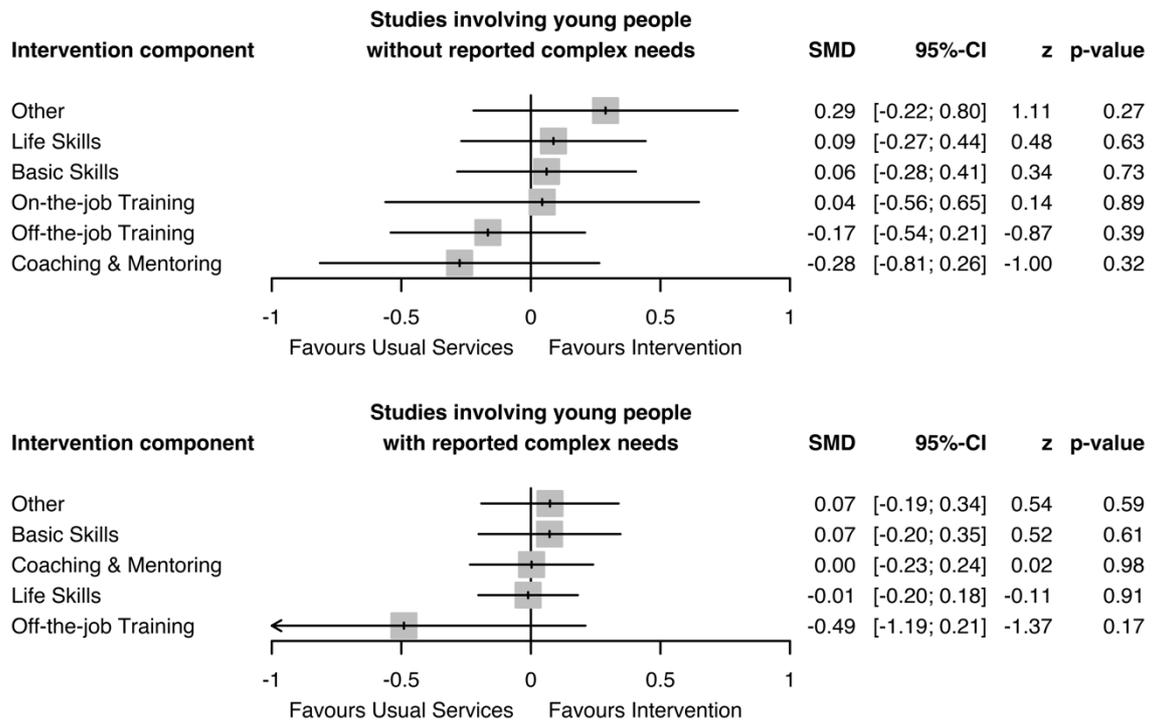


Subgroup analysis of education completion by study population

Thirteen (n=13) of the included studies reporting education completion outcomes served populations who reported facing additional barriers, while the remaining ten (n=10) served populations who did not report facing additional barriers. In this case, there were insufficient studies to undertake a subgroup analysis by study population.

The results — depicted in Figure 25 — show that there are no significant differences in education completion outcomes by study population. Additionally, there are no significant differences in outcomes between populations. The wide confidence intervals depicted in the forest plots indicate both that the sample is underpowered and that our estimates are uncertain.

Figure 25 Forest plot depicting the result of subgroup analysis for education completion by study population with reported additional barriers



Sensitivity analysis

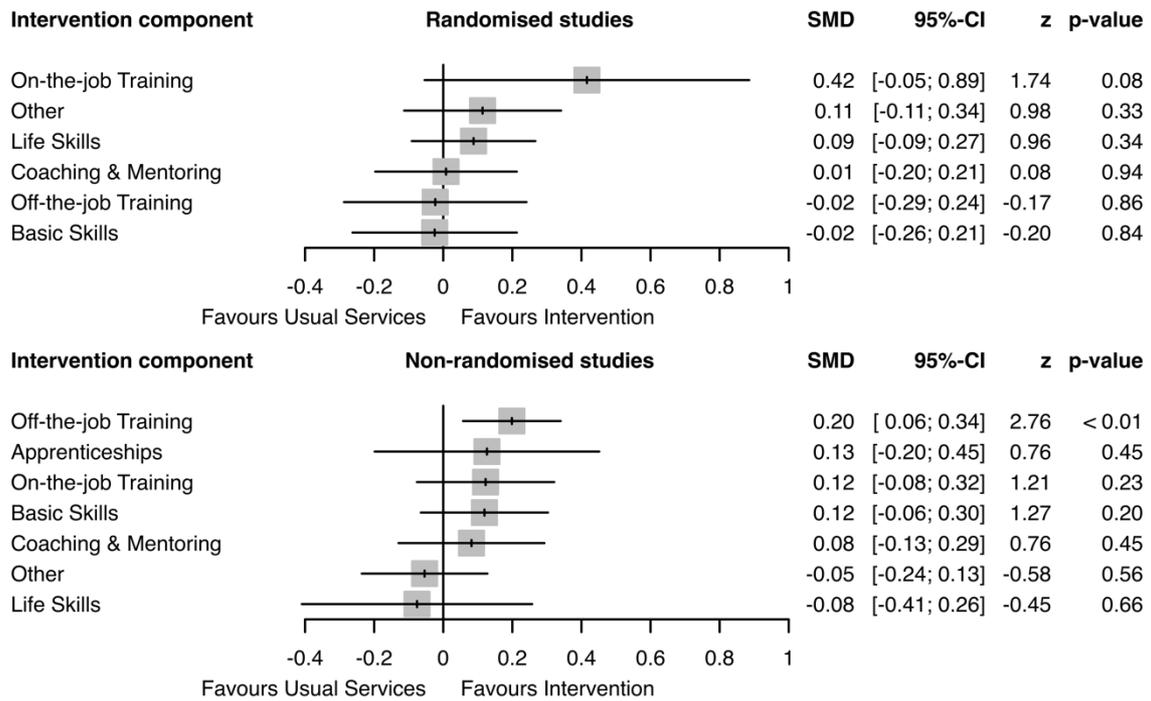
The sensitivity of results was explored by comparing results between studies that used randomised and used non-randomised designs.

Sensitivity analysis of employment status by study design

Thirty (n=30) of the fifty-five (n=55) included studies that report employment status outcomes used randomised study designs, while the remaining twenty-five (n=25) used non-randomised designs.

The results of this analysis — presented below in Figure 26 — suggest that the observed overall positive effect of off-the-job training in the additive CNMA may be driven by the larger observed effect sizes in non-randomised studies. The absence of a statistically significant positive effect among randomised studies may be a result of a lack of statistical power — there are fewer randomised than non-randomised studies — it may also reflect a more accurate estimate of effect (i.e., the true effect is that off-the-job training has no impact on employment status) from studies that use designs which generally provide a higher level of confidence.

Figure 26 Forest plot depicting the result of sensitivity analysis for employment status by study design



Sensitivity analysis of education completion by study design

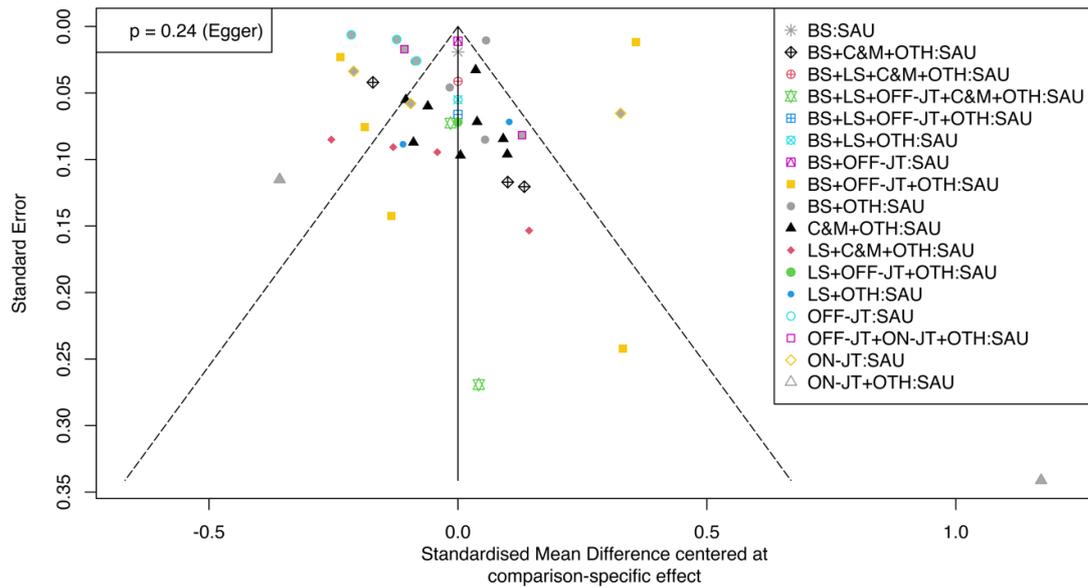
Of the twenty-three (n=23) studies reporting education completion outcomes, all but three (n=20) use randomised designs. This meant that it was not possible to undertake sensitivity analysis by study design for this outcome.

Assessing publication bias

The presence of multiple comparators can make it tricky to apply tools typically used to assess publication bias in pairwise meta-analysis to an NMA. Since the majority of included studies for both outcomes use 'services as usual' as a comparator we think it's appropriate to use a 'comparison-adjusted funnel plot'. A comparison adjusted funnel plot is a modified funnel plot that allows for the comparison of all studies in the network irrespective of the components that they compare (Chaimani & Salanti, 2012). This method can be used to identify possible small-study effects through either visual inspection of the funnel plot and applying Egger's test for funnel plot asymmetry to test it quantitatively (Egger et al., 1997).

The comparison adjusted funnel plot for employment status — depicted in Figure 27 — does not indicate the presence of any asymmetry. This is supported by the findings of Egger's test ($t(45) = 1.19, p > 0.05$). Taken together this indicates that we should not be concerned about the presence of publication bias for this outcome.

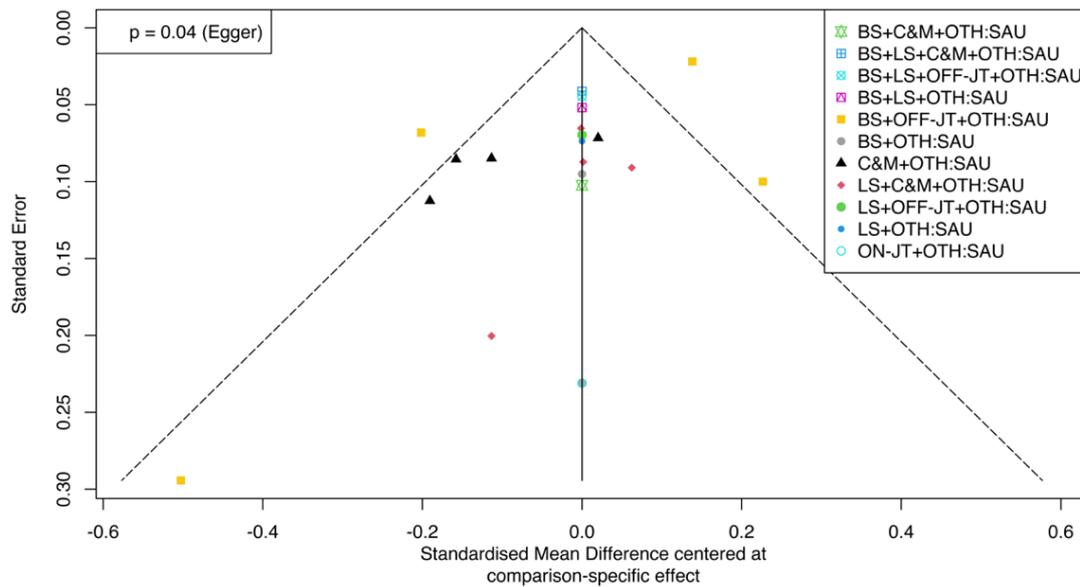
Figure 27 Comparison-adjusted funnel plot for employment status¹²



However, for education completion both the comparison adjusted funnel plot — depicted in Figure 28 — and Egger's test ($t(19) = 0.152, p < 0.05$) indicate the presence of funnel plot asymmetry. This suggests that we should be concerned about the possibility of publication bias that favours the intervention for this outcome.

¹² **Plot legend** — BS: Basic Skills, LS: Life Skills, OFF-JT: Off-the-job training, ON-JT: On-the-job-training, APP: Apprenticeships, C&M: Coaching and mentoring, OTH: Other (residual) component.

Figure 28 Comparison-adjusted funnel plot for education completion¹³



Assessing network coherence

Network coherence was assessed at both the local-level — through the application of node-splitting — and at the global level — through fitting a design-by treatment model.

The results for employment status suggest that there is a moderate amount of incoherence present in the network of included studies that report that outcome. The level of incoherence does not present a major concern to the validity of the results, however it does suggest that some caution should be taken in drawing conclusions from combinations of components that rely solely on indirect comparisons (or a high proportion of mixed evidence). The full results of the analysis of network coherence for the employment status outcome are detailed in Appendix E.

It was not possible to assess global or local incoherence within the network for education completion due to the absence of indirect comparisons within the network for included studies report that outcome.

¹³ **Plot legend** — BS: Basic Skills, LS: Life Skills, OFF-JT: Off-the-job training, ON-JT: On-the-job-training, APP: Apprenticeships, C&M: Coaching and mentoring, OTH: Other (residual) component.

Discussion

This review collated existing evidence that evaluated employment and skills programmes for young people in high income countries. Using a network meta-analysis approach, the impact of individual components that were delivered as part of the evaluated programmes was assessed using two key outcomes: employment, and education completion. The components that were assessed for their effectiveness with respect to these outcomes were: basic skills, life skills, on-the-job training, off-the-job training, apprenticeships and coaching and mentoring.

Summary of key findings

Study characteristics

A considerable number of relevant primary studies ($n=60$) were identified and included in the quantitative synthesis, 32 of these studies used a randomised study design, and 28 a non-randomised quasi-experimental approach. It is noteworthy that only eleven studies were assessed as high confidence (i.e., we can have a high confidence in the study's methodology and findings). A majority of the studies were undertaken in the United States, with far fewer conducted in Europe or the United Kingdom; and two-thirds were published after 2010. Seventeen ($n=17$) studies involved populations where the majority (greater than 50 per cent) reported facing additional barriers.

All of the components of interest to this review were identified in the programmes evaluated in the included studies. Coaching and mentoring ($n=25$) and basic skills ($n=22$) components were the most common components delivered as part of the included programmes. The components were not typically provided in isolation. For example, the Danish programme 'Bridging the Gap between Welfare and Education' provided a range of services which included 'basic skills', 'life skills', 'coaching and mentoring' and 'other' (Rosholm et al., 2019). This is perhaps not surprising given that real world employment and skills programmes usually provide multiple components in combination — this may include components of interest to this review, with or without other programme elements (e.g., case management).

What components of the programmes were effective?

One of the benefits of component network meta-analysis methods is that they can provide an assessment of the relative contribution of each component of a wider programme on outcomes of interest. Our overall analyses provide evidence that a common component of employment and skills programmes — specifically off-the-job training ($g=0.13$, 95% CI: [0.01; 0.25], $p < 0.05$) — had a moderate-sized and statistically significant impact on the employment status for young people who have typically accessed these services. No other components had a statistically significant effect.

Were the components more effective for some subgroups?

Further analyses explored whether the effectiveness of the components differed when separating the data into different subgroups. We assessed the influence of: study design (randomised vs. non-randomised methods); study location (United States vs. other high-income countries); study confidence (high confidence vs. low and medium confidence); and, study participants (general population versus young people identified as facing additional barriers, i.e., those living with a disability, with mental health conditions, or with prior experience of the out-of-home care or juvenile justice systems).

A key result was that the impact of 'on-the-job training' and 'off-the-job training' — one of which had a small significant effect overall — was significantly amplified when provided to young people who report facing additional barriers. For this subgroup, the impact of these components is substantial. Both 'on-the-job training' ($g=1.58$, 95% CI: [0.88, 2.28], $p < 0.01$) and 'off-the-job training' ($g=0.59$, 95% CI: [0.08, 1.11], $p < 0.05$) had high impacts on employment status for young people facing additional barriers. At the same time, off-the-job training had a moderate impact ($g=0.14$, 95% CI: [0.02, 0.27], $p < 0.05$) on employment status for young people who did not face additional barriers. There were no other statistically significant differences between different subgroups.

Interactions between components

We sought to identify whether there were any interactions between different components which might amplify (or nullify) their effect. For example, when 'on-the-job training' and 'life skills' were delivered together in one programme, did this lead to significantly better or worse outcomes than when they were delivered as individual components?

All possible two-way interactions were identified across the studies. That is, where any two components existed in a programme, and there was an opportunity for them to interact. Five two-way interactions were evaluated for their effect on the employment status of young people: a) basic skills x off-the-job training; b) on-the-job training x other; c) basic skills x other; d) life skills x other; and, e) coaching & mentoring x other. None of the interactions had a statistically significant impact on employment outcomes, nor did they affect a significant change in any of the other components. Four two-way interactions were evaluated for their effect on education completion: a) on-the-job training x other, b) basic skills x other, c) life skills x other and d) coaching & mentoring x other. As with employment status, none of the interactions were significant.

Consideration of residual components

The employment and skills programmes evaluated across the included studies incorporated the components of interest to this review. However, more often than not the programs also provided 'something else', i.e., other components that were not the focus of this review. These were grouped into

a residual component group, and six possible components were identified within this group: case management, paid work experience, brokerage and referrals, counselling and programme access.

While further analyses were undertaken to explore this heterogeneity within the residual components group, these have not been reported for multiple reasons. First, by adding additional components we created additional nodes in the network. By doing this we also increased its sparseness to the point that it threatened the coherence of the network (i.e., it had too many separate elements that were not linked). Second, these components were not part of our initial protocol and were not included in the search or screening process and therefore some studies that reported on them may have been missed. This would present a substantial risk of bias.

Recommendations for practice and policy

Young people not in employment, education or training face a range of barriers to securing and maintaining employment. The overall findings from this review suggest there is no panacea for this, however some suggestive recommendations for practice and policy do emerge.

The most substantial finding of this review was that two commonly delivered components — on-the-job training and off-the-job training — have a large effect on employment outcomes for young people who report facing additional barriers. This finding suggests that there may be merit in an approach to commissioning that involves the targeted implementation of these specific programme components for such youth.

At the population-level, the component with the largest treatment effect for employment status was 'off-the-job training'. With this in mind, it could be beneficial for providers to ensure that training provided is of high quality and aligned with current and projected labour market needs.

Based on the current available evidence, we would be hesitant to make recommendations to policymakers and commissioners making decisions about which specific components to include in an employment and skills programme. However, the findings do suggest that on-the-job and off-the-job training may be beneficial inclusions in any programme targeting youth employment outcomes. Given the limitations of the body of evidence reviewed, we were not able to conclude whether other components were effective – therefore they should not be excluded from programmes based on the evidence reported here. Before determining appropriate programme components, we would recommend a) considering the characteristics of the target cohort and what their skills needs might be, b) considering implementation factors such as the required intensity of support and mode of delivery, in order to maximise engagement.

These recommendations should be considered in the context of the overall body of evidence reviewed. While a considerable range of studies were identified, the overall confidence was low to moderate. These studies were largely undertaken in the United States, with far fewer from the United

Kingdom or other countries, with a majority published after 2010. This is notable because the programmes, as well as the components of interest to this review, have been delivered to young people across high-income countries for several decades. This suggests that the studies do not represent what is actually delivered, and that there is a need for these programmes to be evaluated more broadly. For example, despite being a common feature of vocational education and training systems in the United Kingdom, Europe and Australasia, the only evidence this review obtained about the relative effectiveness of apprenticeship programmes on employment outcomes was sourced from two studies undertaken in a single jurisdiction in the United States. That said, overall, the studies did evaluate programmes that included a convincing range of the components of interest which allowed for CNMA analyses.

Using the available evidence, we did not identify statistically significant positive impacts for several of the components commonly delivered in employment and skills programmes. Because we are limited by the available evidence (i.e., that which has been evaluated using high quality methods and published), this finding should not be interpreted as these components are not effective.¹⁴ The number of studies that include each component can also limit the precision with which we could detect an effect by not providing sufficient statistical power. Apprenticeships (n=2) and on-the-job training (n=7) were only included in a small number of studies, yet overall, the results, while not statistically significant, showed a positive direction of effect.

To further develop this evidence base, we would strongly encourage organisations commissioning employment and skills programs for young people to fund and support methodologically rigorous evaluations that use experimental or quasi-experimental methods.

Recommendations for research

There are a range of recommendations for future research that emerge from this review.

Firstly, there is a clear need for more rigorous primary research on the impact of employment and skills programmes in settings outside the United States. This is particularly evident for apprenticeships in high income countries other than the United States and on-the-job training, but this equally applies to other components of employment and skills programmes.

¹⁴ Although we were unable to detect any statistically significant differences between any of the other components on employment status, or any components on education status this does not mean that these components are ineffective. Our choice of method — i.e., CNMA — allows for us to determine if a component of a program has an impact on the outcome of interest. If no difference is detected, it means that we don't have enough evidence to reject the null hypothesis (i.e., that there is no difference between those that received the component and those that did not). The flip side is that, if a difference is detected, then we can be confident that it is present.

In the analysis and reporting of results it would be helpful if future primary research utilised analysis methods that measured the impact of the programme on outcomes over time. Most of the current research relies on post-test only outcomes that do not always control for an individual's characteristics at baseline (e.g., if they were currently or previously employed). Additionally, quantitative treatment effects should be reported as effect sizes or with measures of sample characteristics (i.e., standard error/deviation) to allow future reviewers to transform them into one.

It would be helpful if future primary research included more detailed information on the study population. Age and gender should be considered the bare minimum to report. As discussed previously, any information that allows for segmentation by high-risk populations (e.g., prior involvement in out-of-home care or juvenile justice), demographics (e.g., age, gender, ethnic group), or prior education and/or employment status would provide useful evidence on the impact of components for different groups of people.

Overall, the programmes themselves were not described comprehensively across the studies. Additional information about the content of included programmes (i.e., what exactly do they do, and for how long and how intensely do they do it) and also what services as usual look like in the setting where the programme is being implemented, would be useful for future reviews. Additionally, comparative effectiveness studies and multi-arm experimental trials would assist in the development of more robust network structures for future network meta-analyses.

There is scope to update or expand on this particular review in the future. Before doing so, it might be worthwhile to systematically assess all employment and skills programmes to identify what components are often provided together and use this — along with stakeholder input — to guide the scope of the extension to the next iteration of the NMA. A future review would then be able to undertake a systematic search that specifically searches for all of these 'other' components (e.g., case management, job search assistance) to construct a more robust network. A future review could explore: a) the use of Bayesian hierarchical NMA to assess the impacts of location and variation in services as usual and b) if length of follow-up time affects the results by exploring how results vary in the period following the intervention.

It appears that none of the components of interest to this review have a significant impact on education completion. However, this does not mean that there are not components of employment and skills programmes that are effective at supporting young people to attain their secondary, high school or equivalent qualification. While it was not statistically significant, the treatment effect for the residual 'other' component in the additive CNMA was the largest of across the included components. It would be beneficial to analyse the components of these programmes to identify what they might be and use them as the basis for a future review that examines education

outcomes. There might be specific components that are beneficial for education outcomes (e.g., remediation).

Strengths and limitations of the review findings

Strengths and limitations of CNMA

The component network meta-analysis utilised in this review provides unique insights for programme designers and policymakers that would not be available using other methods (e.g., pairwise meta-analysis or even a standard network analysis that compares combinations of components). The major benefit of a CNMA is its ability to disentangle the relative contribution of each component and assess their effect in combination or alone. This is a major advantage over a standard NMA — where it is only possible to assess the effectiveness of combinations of components as they are delivered (i.e., as part of programmes). As a result, the findings of a CNMA can provide programme designers and policy makers with guidance on what components might be best to implement and test.

Like all methods CNMA relies on some assumptions, violations of which can introduce bias. The major one is the additivity assumption which assumes that the effect of each programme with multiple components can be estimated by summing the relative effect of its included components. This can be a heroic assumption if some components in a network are expected to be reinforcing. While it is possible to account for an interaction between components using an interaction model, data limitations limit their use. In the context of this review, we have assessed that the additivity assumption is reasonable, largely because we have no evidence of any interactions that may violate it.

Another key consideration is the consistency between results estimated from direct evidence (i.e., head-to-head comparisons) and indirect evidence (i.e., that inferred by the network). We identified some incoherence between direct and indirect evidence in the network for employment status.¹⁵ The implication of this incoherence is that some caution should be taken when drawing conclusions from findings that rely on indirect (or high proportions of mixed) evidence i.e., decision makers may want to prioritise the use of direct-evidence. That said, we are confident in the integrity of the network and that, by and large, the mixed and indirect evidence is consistent with the direct evidence (i.e., the conclusions are generally consistent). We are not overly concerned about the potential for bias here as the vast majority of evidence is direct.

Limitations of a rapid review methodology

Ideally, this review would have been undertaken using a systematic review methodology to identify every potentially relevant study on this topic. The overall aim of this analysis was to inform the publication of a practical toolkit,

¹⁵ Detail on the nature of this incoherence is included in Appendix B.

and so time and resource constraints necessitated the use of a rapid review methodology to inform this CNMA. While the review team undertook extensive analyses of the included studies, the fact remains that our search strategy relied heavily on the assumption that the White and Apunyo (2021) EGM was able to identify all of the relevant literature on this topic. While we have confidence in the methods employed by the EGM and their implementation, our inability to validate their results is a limitation. Another limitation is our choice to undertake data extraction by a single reviewer. Whilst an experienced reviewer always double checked the accuracy of this data extraction, this method is not as rigorous as independently double extracted data. While these are well recognised limitations of rapid reviews, they are nonetheless important to highlight.

Potential for limitations in external validity

The components of employment and skills programmes of interest to this review are widely implemented in many settings around the world — particularly in high-income countries. Whilst the review team is confident that the results of this review are based on a reasonably robust summary of the available evidence, it should not be considered the final word on the effectiveness of each of these components. To cite one example, apprenticeships are a key component of vocational education and training in many countries around the world, yet this review was only able to identify two studies that examined the same apprenticeship programme in the same jurisdiction in the United States. Since these two studies are not necessarily representative of apprenticeship programmes that are widely available in high-income countries, it would be inappropriate to conclude that apprenticeships are ineffective based on these findings.

Limitations in considering residual components

In the process of conducting the review it became clear that it was important to consider the role of 'other' components of programmes — beyond those of interest to this review — this may have an impact on outcomes. While we are confident that this was the most appropriate approach to lower potential bias, there are limitations that are important to highlight. In particular, we did not pre-specify that we intended to do this in the protocol. As a result, we did not specifically search for any of these 'other' components, and this may have biased our results as they may not represent the existing evidence of 'other' components. In other words, if we specifically searched for programmes that included 'other' components — such as 'job search assistance' — we may have identified additional programs that could have potentially changed the structure of our network. That said, we believe that considering a 'consolidated other' component is essential for minimising bias associated with other active components in the included programmes. In other words, we controlled for this bias to a point, but we could not describe the individual residual elements or their impact.

Limitations in consideration of follow up time

Following our protocol, we extracted outcomes at the last point at which they were reported. This approach could possibly bias results toward the null hypothesis (i.e., that there is no difference between those who receive the intervention and the comparison group). This is because some components of employment and skills programmes may potentially speed up the process of attaining employment, but over time the comparison condition catches up (Groh et al., 2016; McKenzie, 2017). By using results from the last time point at which they are reported, we could have underestimated the potential for components to promote earlier employment attainment.

Strengths and limitations of the available evidence

Strengths of available evidence

The amount of evidence that was available to inform the analysis of employment status was a strength. Of the sixty (n=60) included studies, fifty-five (n=55) included an outcome measure that was able to be included in a quantitative synthesis. This allowed for subgroup analysis by:

- Study confidence — sixteen (n=16) studies considered to be 'high confidence', whilst the remaining forty-five (n=45) were adjudged to be 'medium or low confidence',
- Populations facing additional barriers — seventeen (n=17) involved populations where the majority of participants with reported facing additional barriers, with the remaining thirty-eight (n=38) involving populations that do not report facing additional barriers,
- Study location — thirty-six (n=36) were conducted in the United States, with the remaining nineteen (n=19) conducted elsewhere.

It was also possible to undertake sensitivity analysis by study design, with thirty (n=30) studies using randomised study designs, and twenty-five (n=25) using non-randomised designs.

Limitations of available evidence

There were limitations in the available evidence that are important to highlight. More than two-thirds of included studies were from the United States — including all of the evidence on the effectiveness of apprenticeships. This suggests that the rest of the world, particularly the United Kingdom and Europe, need to increase their investment into undertaking rigorous evidence of their employment and skills programmes. It was also surprising that no evidence was available for high-income settings in Asia.

There were some exceptions, but many studies did not include sufficient detail about what the programme or intervention does or how it works. This limits what we can say about the interventions by making it hard to assess what participants actually received or even to determine whether it is

suitable (or even possible) to implement in another context. In general, published studies tended to include few details (particularly papers from the economics literature), with more detail reported within the grey literature. Reports from one institution in particular — MDRC, a U.S. social research organisation — are exemplars of the level of detail that would be required to advance knowledge in this area.

Similarly, few studies reported sufficient information about their comparison conditions. Most compared the programme or intervention to ‘services as usual’, but then provided very little detail on what ‘services as usual’ might look like for typical participants. This information is critically important for appraising the heterogeneity of studies and therefore their suitability for quantitative synthesis.

Limitations in the reporting of results

Many included studies did not report quantitative results in sufficient enough detail to allow us to transform the results into a common effect size. In many cases, they did not include standard errors or standard deviations. In some cases, these could be sourced from authors upon request, however this was not the norm. The absence of reporting these basic results prevented the review team from undertaking quantitative synthesis of two additional outcomes for this review.

In almost all of the included studies, outcomes were only reported as ‘post-test only’. This means that we could only determine how many participants had attained the outcome at the time it was last measured, as opposed to how many attained the outcome during the intervention period (i.e., between when they started the intervention and the end of the follow-up period). Post-test only measurements can produce a biased estimate by not accounting for the incidence of the outcome at baseline. However, if the study was a well conducted RCT, or used QED methods that controlled for employment status or education completion at baseline, then it is plausible to assume that this bias is distributed evenly between the two groups.

Finally, there was inconsistent reporting of the demographic characteristics of study populations across the included studies. Some studies provided a comprehensive breakdown of participant characteristics by demographics (e.g., age and gender) and life experiences (e.g., lived experience of mental health condition, living with a disability, prior experience of out-of-home care and/or juvenile justice) that allowed us to analyse different subgroups separately. Other studies provided few, if any, details. This is a limitation because while young people not in employment, education or training are by definition a disadvantaged group, there are some members of this group who face additional barriers — being able to identify which programs are effective and implementable for those young people, is a necessary step to understanding if what is effective for them varies from what is effective for the general population.

Conclusions

The primary goal of this review was to provide input into the first iteration of the Youth Futures Foundation's online evidence toolkit. This toolkit will provide policymakers and programme designers in the United Kingdom with details on the state of the evidence on the relative effectiveness of a range of components that are commonly provided as part of employment and skills programmes. While the review will provide important insights for the first iteration of the toolkit, it is important to consider that the evidence base is dynamic and that this review will need to be repeated to ensure the toolkit remains relevant and accurate.

This review found that on-the-job training can have a statistically significant moderate impact on improving employment outcomes for young people who are not in employment, education or training. The effect of both on-the-job and off-the-job training is hugely amplified when provided to young people who report experiencing additional barriers, suggesting that there is some benefit in targeting these components to particular populations.

There are numerous opportunities for future research to strengthen the evidence base, particularly by undertaking primary research outside the United States. There are also opportunities to repeat and extend the methods used in this review to provide additional insights on the impact of other components of employment and skills programmes.

References

- Alegre, M. À., Casado, D., Sanz, J., & Todeschini, F. A. (2015). The impact of training-intensive labour market policies on labour and educational prospects of NEETs: Evidence from Catalonia (Spain). *Educational Research*, 57(2), 151–167.
<https://doi.org/10.1080/00131881.2015.1030852>
- Balduzzi, S., Rücker, G., Nikolakopoulou, A., Papakonstantinou, T., Salanti, G., Efthimiou, O., & Schwarzer, G. (2023). netmeta: An R Package for Network Meta-Analysis Using Frequentist Methods. *Journal of Statistical Software*, 106(2). <https://doi.org/10.18637/jss.v106.i02>
- Bampasidou, M. (2012). *Unbundling the degree effect in a job training program for disadvantaged youths* [University of Miami].
<https://ufdc.ufl.edu/UFE0044581/00001/pdf>
- Bampasidou, M., Flores, C. A., Flores-Lagunes, A., & Parisian, D. J. (2014). The role of degree attainment in the differential impact of job corps on adolescents and young adults. *Research in Labor Economics*, 40, 113–156. <https://doi.org/10.1108/S0147-912120140000040004>
- Bauer, E. L., Crosse, S., McPherson, K., Friedman, J., Zacharia, Z., Tapper, D., & Clarke, R. (2014). *Evaluation of the New York City Justice Corps: Final Outcome Report* (p. 179). WeStat and Prisoner Reentry Institute, John Jay College of Criminal Justice.
<https://www1.nyc.gov/assets/opportunity/pdf/Westat-Justice-Corps-Evaluation.pdf>
- Blanco, G., Flores, C. A., & Flores-Lagunes, A. (2013a). Bounds on average and quantile treatment effects of job corps training on wages. *Journal of Human Resources*, 48(3), 659–701. <https://doi.org/10.3368/jhr.48.3.659>
- Blanco, G., Flores, C. A., & Flores-Lagunes, A. (2013b). The Effects of Job Corps Training on Wages of Adolescents and Young Adults. *American Economic Review: Papers & Proceedings*, 103(3), 418–422.
<https://doi.org/10.1257/aer.103.3.418>
- Blanco, G., & Flores-Lagunes, A. (2017). *Does Youth Training Lead to Better Job Quality? Evidence from Job Corps*. 45.
- Bloom, H. S., Orr, L. L., Bell, S. H., Cave, G., Doolittle, F., Lin, W., Bos, J. M., Bloom, H. S., Orr, L. L., Bell, S. H., & Doolittle, F. (1997). The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study. *The Journal of Human Resources*, 32(3), 549–576.

- Bloom, H. S., Orr, L. L., Cave, G., Bell, S. H., & Doolittle, F. (1993). *The National JTPA Study: Title II-A Impacts on Earnings and Employment at 18 Months* (p. 458). Abt Associates.
- Blundell, R., Dias, M. C., Meghir, C., & van Reenen, J. (2004). Evaluating the Employment Impact of a Mandatory Job Search Program. *Journal of the European Economic Association*, 2(4), 569–606. <https://doi.org/10.1162/1542476041423368>
- Borenstein, M., Higgins, J. P. T., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity: I^2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, 8(1), 5–18. <https://doi.org/10.1002/jrsm.1230>
- Brunetti, I., & Corsini, L. (2017). Workplace training programs: Instruments for human capital improvements or screening devices? *Education + Training*, 59(1), 31–46. <https://doi.org/10.1108/ET-09-2014-0104>
- Buchanan, I. (2023). *Youth Unemployment Statistics*. House of Commons Library. <https://researchbriefings.files.parliament.uk/documents/SN05871/SN05871.pdf>
- Caliendo, M., Künn, S., & Schmidl, R. (2011). *Fighting Youth Unemployment: The Effects of Active Labor Market Policies*. IZA. <https://doi.org/10.2139/ssrn.1977808>
- Canzian, G., Foubert, J., Meroni, E. C., & Santangelo, G. (2020). *Active Labour Market Policies in Flanders. Evaluation of the ESF “Work Experience for Young Persons” programme* (EUR 30144 EN). Publications Office of the European Union. doi:10.2760/623819
- Cave, G., Bos, H., Doolittle, F., & Toussaint, C. (1993). *JOBSTART: Final Report on a Program for School Dropouts*. Manpower Demonstration Research Corporation. <https://www.ojp.gov/ncjrs/virtual-library/abstracts/jobstart-final-report-program-school-dropouts#:~:text=This document is the final report on the,the employment and earning prospects for school dropouts.>
- Centeno, L., Centeno, M., & Novo, A. A. (2008). *Evaluating Job Search Programs for Old and Young Individuals: Heterogeneous Impact on Unemployment duration* [Estudos e Documentos de Trabalho]. Banco de Portugal Eurosistema.
- Centeno, M., & Novo, A. A. (2006). Avaliação Do Impacto De Um Programa De Apoio À Procura De Emprego. *Boletim Económico Do Banco de Portugal, Verão*, 85–96.

- Chaimani, A., Caldwell, D. M., Li, T., Higgins, J. P. T., & Salanti, G. (2022). Chapter 11: Undertaking network meta-analyses. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions: Vol. Version 6.3* (updated February 2022). Cochrane. <https://training.cochrane.org/handbook>
- Chaimani, A., & Salanti, G. (2012). Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. *Research Synthesis Methods, 3*(2), 161–176. <https://doi.org/10.1002/jrsm.57>
- Chen, X. (2013). *Partial Identification of Average Treatment Effects in Program Evaluation: Theory and Applications* [University of Miami]. <http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=1471063&site=ehost-live>
- Chen, X., Flores, C. A., & Flores-Lagunes, A. (2018). Going beyond LATE: Bounding Average Treatment Effects of Job Corps Training. *Journal of Human Resources, 53*(4), 1050–1099. <https://doi.org/10.3368/jhr.53.4.1015.7483R1>
- Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences* (1st Edition). Routledge Academic.
- Courtney, M. E., Valentine, E. J., & Skemer, M. (2019). Experimental evaluation of transitional living services for system-involved youth: Implications for policy and practice. *Children and Youth Services Review, 96*(June 2018), 396–408. <https://doi.org/10.1016/j.childyouth.2018.11.031>
- Courtney, M. E., Zinn, A., Johnson, H., & Malm, K. E. (2011). *Evaluation of the Massachusetts Adolescent Outreach Program for Youths in Intensive Foster Care: Final Report* (OPRE Report #2011-14). Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. <https://www.acf.hhs.gov/opre/report/evaluation-massachusetts-adolescent-outreach-program-youths-intensive-foster-care-final>
- Courtney, M. E., Zinn, A., Koralek, A., & Bess, R. J. (2011). *Evaluation of the Independent Living – Employment Services Program, Kern County, California: Final Report* (OPRE Report #2011-13). Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. <https://www.acf.hhs.gov/opre/resource/evaluation-of-the-independent-living-employment-services-program-kern>
- Cummings, D., Farrell, M., & Skemer, M. (2018). *Forging a Path: Final Impacts and Costs of New York City’s Young Adult Internship Program* (OPRE

- Report 2018-75). OPRE Report 2018-75. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Davis, J. M. V., & Heller, S. B. (2017). *Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs* (Working Paper 23443). National Bureau of Economic Research. <http://www.nber.org/papers/w23443>
- De Giorgi, G. (2005). The New Deal for Young People Five Years On. *Fiscal Studies*, 26(3), 371–383.
- Department for Work & Pensions,. (2023). *Official Statistics: Employment of disabled people 2022*. GOV.UK. <https://www.gov.uk/government/statistics/the-employment-of-disabled-people-2022/employment-of-disabled-people-2022>
- Donato, L., Migliore, M. C., & Poy, S. (2018). Employment Effects of Vocational Training: An Evaluation Using Propensity Score Matching. *Politica Economica*, 34(3), 273–296. <https://doi.org/10.1429/92121>
- Duarte, N., Geraci, A., Granato, S., Mazzarella, G., & Mortagua, M. J. (2020). *The evaluation of the Youth Employment Initiative in Portugal using Counterfactual Impact Evaluation methods* (JRC Technical Report). Joint Research Centre. <https://publications.jrc.ec.europa.eu/repository/handle/JRC120942>
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629. <https://doi.org/10.1136/bmj.315.7109.629>
- Ehlert, C. R., Kluve, J., & Schaffner, S. (2012a). Temporary Work as an Active Labor Market Policy: Evaluating an Innovative Activation Program for Disadvantaged Youths. *Economics Bulletin*, 32(2), 1765–1773.
- Ehlert, C. R., Kluve, J., & Schaffner, S. (2012b). *Temporary Work as an Active Labor Market Policy: Evaluating an Innovative Program for Disadvantaged Youths*. Institute for the Study of Labor. <https://doi.org/10.2139/ssrn.2096800>
- Fein, D., & Hamadyk, J. (2018). *Bridging the Opportunity Divide for Low-Income Youth: Implementation and Early Impacts of the Year Up Program* (OPRE Report #2018-65). Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Flores-Lagunes, A., Gonzalez, A., & Neumann, T. (2010). Learning but not earning? The impact of job corps training on hispanic youth. *Economic Inquiry*, 48(3), 651–667. <https://doi.org/10.1111/j.1465-7295.2009.00211.x>

- Fraker, T., Baird, P., Mamun, A., Manno, M., Martinez, J., Reed, D., & Thompkins, A. (2012). *The Social Security Administration's Youth Transition Demonstration Projects: Interim Report on the Career Transition Program*. Mathematica Policy Research. <https://www.mathematica.org/publications/the-social-security-administrations-youth-transition-demonstration-projects-interim-report-on-the-career-transition-program>
- Fraker, T. M., Baird, P., Black, A., Mamun, A. A., Manno, M., Martinez, J., Rangarajan, A., & Reed, D. (2011). *The Social Security Administration's Youth Transition Demonstration Projects: Interim Report on Colorado Youth Wins*. Mathematica Policy Research. <https://www.mathematica.org/publications/the-social-security-administrations-youth-transition-demonstration-projects-interim-report-on-colorado-youth-wins>
- Fraker, T. M., Black, A., Broadus, J., Mamun, A., Manno, M., Martinez, J., McRoberts, R., Rangarajan, A., & Reed, D. (2011). *The Social Security Administration's Youth Transition Demonstration Projects Interim Report on the City University of New York's Project*. Mathematica Policy Research. <https://www.mathematica.org/publications/the-social-security-administrations-youth-transition-demonstration-projects-interim-report-on-the-city-university-of-new-yorks-project>
- Fraker, T. M., Black, A., Mamun, A., Manno, M., Martinez, J., O'Day, B., O'Toole, M., Rangarajan, A., & Reed, D. (2011). *The Social Security Administration's Youth Transition Demonstration Projects: Interim Report on Transition WORKS* (Mathematica Reference Number: 06209.125). Mathematica Policy Research. <https://www.mathematica.org/publications/the-social-security-administrations-youth-transition-demonstration-projects-interim-report-on-transition-works>
- Fraker, T. M., Cobb, J., Hemmeter, J., Luecking, R., & Mamun, A. (2018). Three-Year Effects of the Youth Transition Demonstration Projects. *Social Security Bulletin*, 78(3), 19–41.
- Fraker, T. M., Crane, K. T., Honeycutt, T. C., Luecking, R. G., Mamun, A. A., & O'Day, B. L. (2018). The youth transition demonstration project in Miami, Florida: Design, implementation, and three-year impacts. *Journal of Vocational Rehabilitation*, 48(1), 79–91. <https://doi.org/10.3233/JVR-170917>
- Fraker, T. M., Honeycutt, T., Mamun, A., Manno, M., Martinez, J., O'Day, B., Reed, D., & Thompkins, A. (2012). *The Social Security Administration's Youth Transition Demonstration Projects: Interim Report on Broadened Horizons, Brighter Futures* (p. 194). Mathematica Policy Research.

- Fraker, T. M., Mamun, A., Honeycutt, T., Thompkins, A., & Valentine, E. J. (2014). *Final Report on the Youth Transition Demonstration Evaluation* (p. 251). Office of Research, Demonstration, and Employment Support, Social Security Administration.
https://www.mdrc.org/sites/default/files/Final_Report_on_the_Youth_Transition_Demonstration_Evaluation.pdf
- Fraker, T. M., Mamun, A., Manno, M., Martinez, J., Reed, D., Thompkins, A., & Wittenburg, D. (2012). *The Social Security Administration's Youth Transition Demonstration Projects: Interim Report on West Virginia Youth Works*. Mathematica Policy Research.
<https://www.mathematica.org/publications/the-social-security-administrations-youth-transition-demonstration-projects-interim-report-on-west-virginia-youth-works>
- Frumento, P., Mealli, F., Pacini, B., & Rubin, D. B. (2012). Evaluating the effect of training on wages in the presence of noncompliance, nonemployment, and missing outcome data. *Journal of the American Statistical Association*, 107(498), 450–466.
<https://doi.org/10.1080/01621459.2011.643719>
- Geckeler, C., Betesh, H., Diaz, H., Folsom, L., Kim, H., & Paprocki, A. (2017). *Helping Dropout Youth Find Education and Employment: Final Impact Report for the Evaluation of the Los Angeles Reconnections Career Academy (LARCA) Program* (p. 153). Social Policy Research Associates.
<https://strategies.workforcegps.org/resources/2019/12/18/21/05/Impact-Report-for-the-Evaluation-of-the-Los-Angeles-Reconnections-Career-Academy-LARCA-Program>
- Gritz, R. M., & Johnson, T. (2001). *National Job Corps Study: Assessing Program Effects on Earnings for Students Achieving Key Program Milestones*. Office of Policy and Research, Employment and Training Administration, U.S. Department of Labor.
<https://wdr.doleta.gov/opr/fulltext/MilestoneImpactReport-Final.pdf>
- Groh, M., Krishnan, N., McKenzie, D., & Vishwanath, T. (2016). Do Wage Subsidies Provide a Stepping-Stone to Employment for Recent College Graduates? Evidence from a Randomized Experiment in Jordan. *Review of Economics and Statistics*, 98(3), 488–502.
https://doi.org/10.1162/REST_a_00584
- Gupta, S., Srinivasan, M., Chen, Y., Patterson, L., & Griffith, T. (2016). *Evaluation of the Linking Innovation, Knowledge, and Employment Program: Final Evaluation Report* (p. 113). IMPAQ International LLC.
<https://strategies.workforcegps.org/resources/2018/05/19/00/12/Evaluation-of-the-Linking-Innovation-Knowledge-and-Employment-Program-Final-Evaluation-Report>

- Hämäläinen, K., & Tuomala, J. (2008). *Vocational Labour Market Training in Promoting Youth Employment*. Government Institute for Economic Research.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Review of Economic Studies*, 64(4), 605–654. <https://doi.org/10.2307/2971733>
- Heckman, J. J., & Smith, J. A. (1999). The pre-programme earnings dip and the determinants of participation in a social programme. Implications for simple programme evaluation strategies. *Economic Journal*, 109(457), 313–348. <https://doi.org/10.1111/1468-0297.00451>
- Higgins, J., Jackson, D., Barrett, J., Lu, G., Ades, A., & White, I. (2012). Consistency and inconsistency in network meta-analysis: Concepts and models for multi-arm studies. *Research Synthesis Methods*, 3(2), 98–110. <https://doi.org/10.1002/jrsm.1044>
- Hollenbeck, K., & Huang, W.-J. (2006). *Net Impact Estimates of the Workforce Development System in Washington State*. Upjohn Institute for Employment Research. <https://doi.org/10.17848/tr06-020>
- Hollenbeck, K., & Huang, W.-J. (2016). *Net Impact and Benefit-Cost Estimates of the Workforce Development System in Washington State* (Upjohn Institute Technical Report 16-033). W.E. Upjohn Institute. <https://doi.org/10.17848/tr16-033>
- International Initiative for Impact Evaluation. (2017). *Youth Employment Evidence Gap Map*. <https://gapmaps.3ieimpact.org/evidence-maps/youth-employment-evidence-gap-map>
- International Labour Organization. (2022). *Global employment trends for youth 2022: Investing in transforming futures for young people*. ILO. <https://doi.org/10.54394/QSMU1809>
- Izzo, M. V., Cartledge, G., Miller, L., Growick, B., & Rutkowsi, S. (2000). Increasing Employment Earnings: Extended Transition Services that Make a Difference. *Career Development for Exceptional Individuals*, 23(2), 139–156.
- Jastrzab, J., Masker, J., Blomquist, J., & Orr, L. (1996). *Impacts of Service: Final Report on the Evaluation of American Conservation and Youth Service Corps*. Corporation for National Service. <https://eric.ed.gov/?id=ED400420>
- Kim, Y., Ju, E., Rosenberg, R., & Farmer, E. (Betsy) M. Z. (2019). Estimating the effects of independent living services on educational attainment and employment of foster care youth. *Children and Youth Services Review*,

96(August 2018), 294–301.
<https://doi.org/10.1016/j.childyouth.2018.11.048>

Kluve, J., Puerto, S., Robalino, D., Romero, J. M., Rother, F., Stöterau, J., Weidenkaff, F., & Witte, M. (2017). Interventions to improve the labour market outcomes of youth: A systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions. *Campbell Systematic Reviews*, 13(1), 1–288.
<https://doi.org/10.4073/csr.2017.12>

Kopečná, V. (2016). Counterfactual Impact Evaluation of the Project Internships for Young Job Seekers. *Central European Journal of Public Policy*, 10(2), 48–66. <https://doi.org/10.1515/cejpp-2016-0026>

Kornfeld, R., & Bloom, H. S. (1999). Measuring program impacts on earnings and employment: Do unemployment insurance wage reports from employers agree with surveys of individuals? *Journal of Labor Economics*, 17(1), 168–197. <https://doi.org/10.1086/209917>

Krahn, U., Binder, H., & König, J. (2013). A graphical tool for locating inconsistency in network meta-analyses. *BMC Medical Research Methodology*, 13(1), 35. <https://doi.org/10.1186/1471-2288-13-35>

Larsson, L. (2003). Evaluation of Swedish youth labor market programs. *Journal of Human Resources*, 38(4), 891–927. <https://doi.org/10.2307/1558784>

Learning and Work Institute. (2022). *Ethnic minority youth employment data analysis*. National Learning and Work Institute.
<https://youthfuturesfoundation.org/wp-content/uploads/2022/10/Final-report-Ethnic-minority-youth-employment-data-analysis.pdf>

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3), 1071–1102. <https://doi.org/10.1111/j.1467-937X.2009.00536.x>

Li, Y., & Heath, A. (2020). Persisting disadvantages: A study of labour market dynamics of ethnic unemployment and earnings in the UK (2009–2015). *Journal of Ethnic and Migration Studies*, 46(5), 857–878.
<https://doi.org/10.1080/1369183X.2018.1539241>

Longhi, S. (2020). A longitudinal analysis of ethnic unemployment differentials in the UK. *Journal of Ethnic and Migration Studies*, 46(5), 879–892.
<https://doi.org/10.1080/1369183X.2018.1539254>

Lüdecke, D. (2019). *esc: Effect Size Computation for Meta Analysis (0.5.1)*.
<https://CRAN.R-project.org/package=esc>

- Maibom, J., Rosholm, M., & Svarer, M. (2014). *Can Active Labour Market Policies Combat Youth Unemployment?* Institute of Labor Economics. <https://docs.iza.org/dp7912.pdf>
- McClanahan, W. S., Sipe, C. L., & Smith, T. J. (2004). *Enriching Summer Work: An Evaluation of the Summer Career Exploration Program*. Public/Private Ventures. <https://eric.ed.gov/?id=ED503188>
- McKenzie, D. (2017). How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence. *The World Bank Research Observer*, 32(2), 127–154. <https://doi.org/10.1093/wbro/lkx001>
- Millenky, M., Bloom, D., & Dillon, C. (2010). *Making the Transition: Interim Results of the National Guard Youth ChalleNGe Evaluation*. MDRC. https://www.mdrc.org/sites/default/files/full_434.pdf
- Millenky, M., Bloom, D., Muller-Ravett, S., & Broadus, J. (2011). *Staying on Course: Three-Year Results of the National Guard Youth ChalleNGe Evaluation*. MDRC. <https://ssrn.com/abstract=2019770>
- Millenky, M., Schwartz, S. E. O., & Rhodes, J. E. (2014). Supporting the Transition to Adulthood among High School Dropouts: An Impact Study of the National Guard Youth Challenge Program. *Prevention Science*, 15(4), 448–459. <https://doi.org/10.1007/s11121-013-0388-4>
- Miller, C., Bos, J. M., Porter, K. E., Tseng, F. M., & Abe, Y. (2005). *The Challenge of Repeating Success in a Changing World: Final Report on the Center for Employment Training Replication Sites* (Issue September, p. 200). MDRC. <http://search.proquest.com/docview/62011515?accountid=13042>
- Miller, C., Cummings, D., Millenky, M., Wiegand, A., & Long, D. (2018). *Laying a Foundation: Four-year Results from the National YouthBuild Evaluation* (p. 155). MDRC. https://www.mdrc.org/sites/default/files/YouthBuild_Final_508%20compliant.pdf
- Miller, C., Millenky, M., Schwartz, L., Goble, L., & Stein, J. (2016). *Building a Future: Interim Impact Findings From the YouthBuild Evaluation* (p. 144). MDRC. https://www.mdrc.org/sites/default/files/YouthBuild_Interim_Report_2016_508.pdf
- Muñoz-Repiso, J. M. C., & Braza, A. S. (2011). Effectiveness of Public Training Programs Reducing the Time Needed to Find a Job. *Estudios de Economía Aplicada*, 29(1), 1–26. <https://doi.org/10.25115/eea.v29i1.3946>

- Murphy, L. (2022). *Not working: Exploring changing trends in youth worklessness in the UK, from the 1990s to the Covid-19 pandemic*. Resolution Foundation. <https://www.resolutionfoundation.org/app/uploads/2022/06/Not-working.pdf>
- Nadon, M. L. (2020). Making the transition: How asset building services can promote positive adult outcomes for foster youth. *Children and Youth Services Review*, 115, 105083. <https://doi.org/10.1016/j.childyouth.2020.105083>
- OECD. (2021). *Designing active labour market policies for the recovery*. (pp. 1–13). Organization for Economic Cooperation and Development. <https://www.oecd.org/coronavirus/policy-responses/designing-active-labour-market-policies-for-the-recovery-79c833cf/>
- Office for National Statistics,. (2023). *X02 Regional labour market: Estimates of unemployment by age*. <https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/datasets/regionalunemploymentbyage02>
- Ott, E., Taylor, D., Rowland, J., Nancarrow, A. F., Newton, B., Featherston, R., & Shlonsky, A. (2022). *Protocol for a network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries*. OSF. <https://doi.org/10.17605/OSF.IO/J63QT>
- Pastore, F., & Pompili, M. (2019). *Assessing the impact of off-and on-the-job training on employment outcomes. A counterfactual evaluation of the PIPOL program*. Global Labour Organization. <http://hdl.handle.net/10419/193688>
- PlotDigitizer Online App. (n.d.). Retrieved November 30, 2022, from <https://plotdigitizer.com/app>
- Price, C., Williams, J., Simpson, L., Jastrzab, J., & Markovitz, C. (2011). *National Evaluation of Youth Corps: Findings at Follow-up*. Corporation for National and Community Service. <https://www.abtassociates.com/insights/publications/report/national-evaluation-of-youth-corps-findings-at-follow-up>
- Puerto, S., Curcio, C., Bausch, J., Stöterau, J., & Weber, M. (2022). *The impact of active labour market programmes on youth: An updated systematic review and meta-analysis*. International Labour Organisation and World Bank. <https://english.iob-evaluatie.nl/results/publications/reports/2022/12/08/evaluation-youth-unemployment>

- Quint, J. C., Bos, Johannes, M., & Polit, D. F. (1997). *New Chance. Final Report on a Comprehensive Program for Young Mothers in Poverty and Their Children*. Manpower Demonstration Research Corporation.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.r-project.org/>
- Roder, A., & Elliott, M. (2011). *A Promising Start: Year Up's Initial Impacts on Low-Income Young Adults' Careers*. Economic Mobility Corporation. https://regroup-production.s3.amazonaws.com/documents/ReviewReference/547117302/25551.pdf?response-content-type=application%2Fpdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAYSFKCAWYQ4D5IUHG%2F20221121%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20221121T010402Z&X-Amz-Expires=604800&X-Amz-SignedHeaders=host&X-Amz-Signature=db8aebfaf6dbb49e0c8246ef507f7f310704e13d52f17139ffdf99af14056de3
- Roder, A., & Elliott, M. (2014). *Sustained Gains: Year Up's Continued Impact on Young Adults' Earnings*. Economic Mobility Corporation.
- Rosholm, M., Mikkelsen, M. B., & Svarer, M. (2019). Bridging the gap from welfare to education: Propensity score matching evaluation of a bridging intervention. *PLOS ONE*, *14*(5), e0216200. <https://doi.org/10.1371/journal.pone.0216200>
- Rotar, L. J. (2012a). Evaluating the Effectiveness of an Institutional Training Program in Slovenia: A Comparison of Methods. *South East European Journal of Economics and Business*, *7*(1), 43–51. <https://doi.org/10.2478/v10033-012-0004-8>
- Rotar, L. J. (2012b). How Effective Is the Slovenian Institutional Training Program in Improving Youth's Chances of Reemployment? *Eastern European Economics*, *50*(3), 94–106. <https://doi.org/10.2753/EEE0012-8775500305>
- Schaeffer, C. M., Henggeler, S. W., Ford, J. D., Mann, M., Chang, R., & Chapman, J. E. (2014). RCT of a promising vocational/employment program for high-risk juvenile offenders. *Journal of Substance Abuse Treatment*, *46*(2), 134–143. <https://doi.org/10.1016/j.jsat.2013.06.012>
- Schochet, P. Z., Burghardt, J., & Glazerman, S. (2001). *National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes*. Office of Policy and Research, Employment and Training Administration, U.S. Department of Labor. https://wdr.doleta.gov/research/FullText_Documents/01-jcimpacts.pdf

- Schochet, P. Z., Burghardt, J., & McConnell, S. (2006). *National Job Corps Study and Longer-Term Follow-Up Study: Impact and Benefit-Cost Findings Using Survey and Summary Earnings Records Data*. Office of Policy Development and Research: Research and Evaluations, Employment and Training Administration, U.S. Department of Labor. https://wdr.doleta.gov/research/FullText_Documents/National%20Job%20Corps%20Study%20and%20Longer%20Term%20Follow-Up%20Study%20-%20Final%20Report.pdf
- Schochet, P. Z., Burghardt, J., & McConnell, S. (2008). Does Job Corps work? Impact findings from the national Job Corps study. *American Economic Review*, 98(5), 1864–1886. <https://doi.org/10.1257/aer.98.5.1864>
- Seitidis, G., Tsokani, S., Christogiannis, C., Kontouli, K.-M., Fyraridis, A., Nikolakopoulos, S., Veroniki, A. A., & Mavridis, D. (2023). Graphical tools for visualizing the results of network meta-analysis of multicomponent interventions. *Research Synthesis Methods*, 14(3), 382–395. <https://doi.org/10.1002/jrsm.1617>
- Skemer, M., Sherman, A., & Williams, S. (2017). *Reengaging New York City's Disconnected Youth Through Work: Implementation and Early Impacts of the Young Adult Internship Program* (OPRE Report 2017-22; p. 140). Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Skemer, M., & Valentine, E. J. (2016). *Striving for independence: Two-year impact findings from the Youth Villages transitional living evaluation*. MDRC. https://www.mdrc.org/sites/default/files/Youth%20Villages_2016_FR.pdf
- Stromback, T. (2010). *Earnings, Schooling and Vocational Education and Training*. 13(3), 23.
- Theodos, B., Pergamit, M. R., Hanson, D., Edelstein, S., & Daniels, R. (2016). *Embarking on College and Career: Interim Evaluation of Urban Alliance* (p. 142). Urban Institute. <https://www.urban.org/research/publication/embarking-college-and-career-interim-evaluation-urban-alliance>
- Theodos, B., Pergamit, M. R., Hanson, D., Edelstein, S., Daniels, R., & Srinivasan, T. (2017). *Pathways after High School: Evaluation of the Urban Alliance High School Internship Program* (p. 206). Urban Institute. <https://www.urban.org/research/publication/pathways-after-high-school-evaluation-urban-alliance-high-school-internship-program>

- Tsokani, S., Seitidis, G., & Mavridis, D. (2022). Component network meta-analysis in a nutshell. *BMJ Evidence-Based Medicine*, bmjebm-2021-111906. <https://doi.org/10.1136/bmjebm-2021-111906>
- United States Administration for Children & Families,. (2022). *Pathways to Work Evidence Clearinghouse*. <https://pathwaystowork.acf.hhs.gov>
- United States Department of Labor,. (2022). *Clearinghouse for Labor Evaluation and Research*. <https://clear.dol.gov>
- Valentine, E. J., Skemer, V. M., & Courtney, M. E. (2015). *Becoming Adults: One-year impact findings from the Youth Villages Transitional Living Evaluation*. MDRC.
- Wasserman, K., Walter, J., Luczywek, B., Wagner, H., & Redcross, C. (2019). *Engaging Young Men Involved in Chicago's Justice System: A Feasibility Study of the Bridges to Pathways Program* (OPRE Report 2019-79). Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. <https://files.eric.ed.gov/fulltext/ED601141.pdf>
- Wehman, P. H., Schall, C. M., McDonough, J., Graham, C., Brooke, V., Riehle, J. E., Brooke, A., Ham, W., Lau, S., Allen, J., & Avellone, L. (2017). Effects of an employer-based intervention on employment outcomes for youth with significant support needs due to autism. *Autism*, 21(3), 276–290. <https://doi.org/10.1177/1362361316635826>
- Wehman, P. H., Schall, C. M., McDonough, J., Kregel, J., Brooke, V., Molinelli, A., Ham, W., Graham, C. W., Erin Riehle, J., Collins, H. T., & Thiss, W. (2014). Competitive Employment for Youth with Autism Spectrum Disorders: Early Results from a Randomized Clinical Trial. *Journal of Autism and Developmental Disorders*, 44(3), 487–500. <https://doi.org/10.1007/s10803-013-1892-x>
- White, H., & Apunyo, R. (2021). *The effectiveness of interventions to increase youth employment: An evidence and gap map*. Youth Futures Foundation.
- White, H., Saran, A., Verma, A., Oprea, E., & Babudu, P. (2022). *Evidence and Gap Map of Interventions to Prevent Children Getting Involved in Violence: Technical Report on the First Edition*. Youth Endowment Fund & Campbell Collaboration. <https://youthendowmentfund.org.uk/wp-content/uploads/2021/02/YEF-Evidence-and-Gap-Map-Technical-Report-FINAL.pdf>
- Wilson, D. B. (n.d.). *Practical Meta-Analysis Effect Size Calculator*. Retrieved September 14, 2020, from <https://campbellcollaboration.org/research-resources/effect-size-calculator.html>

- World Bank. (2022). *World Bank Country and Lending Groups*.
<https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- Zhang, J. L., Rubin, D. B., & Mealli, F. (2009). Likelihood-based analysis of causal effects of job-training programs using principal stratification. *Journal of the American Statistical Association*, 104(485), 166–176.
<https://doi.org/10.1198/jasa.2009.0012>
- Zinn, A., & Courtney, M. (2017). Helping foster youth find a job: A random-assignment evaluation of an employment assistance programme for emancipating youth. *Child and Family Social Work*, 22(1), 155–164.
<https://doi.org/10.1111/cfs.12212>

Appendix A Additional detail on the characteristics of included studies

Table 8 Characteristics of Included Studies

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #1: Alegre et al. (2015)	PQPI	<ul style="list-style-type: none"> Off-the-job training On-the-job training Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Spain <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> not reported (Intervention: n=1220; Comparison: not reported)¹⁶</p>
Study #2: Bauer et al. (2014)	New York City Justice Corps	<ul style="list-style-type: none"> Basic Skills Coaching & Mentoring Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings High school (or equiv.) completion Vocational Education commencement 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Medium <i>Sample size:</i> n=553 (Intervention: n=291; Comparison: n=242)</p>
Study #3: Bloom et al (1993)	Job Training Partnership Act (JTPA) — Classroom training	<ul style="list-style-type: none"> Basic Skills Off-the-job training Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Medium <i>Sample size:</i> n=1571¹⁷ (Intervention: not reported; Comparison: not reported)</p>

¹⁶ Control group sizes estimated — see Table 9

¹⁷ Treatment and control group sizes estimated — see Table 9

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #4: Bloom et al (1993)	Job Training Partnership Act (JTPA) — OJT/JSA	<ul style="list-style-type: none"> On-the-job training Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Medium <i>Sample size:</i> n=1160 (Intervention: not reported; Comparison: not reported)¹⁸</p>
Study #5: Bloom et al (1993)	Job Training Partnership Act (JTPA) — Other services	<ul style="list-style-type: none"> Basic Skills Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=1317 (Intervention: not reported; Comparison: not reported)¹⁹</p>
Study #6: Brunetti & Corsini (2017)	Workplace Training Programs	<ul style="list-style-type: none"> On-the-job training 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status 	<p><i>Design:</i> Randomised <i>Location:</i> Italy <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=4087 (Intervention: n=252; Comparison: n=3835)</p>
Study #7: Centeno et al. (2008)	Inserjovem	<ul style="list-style-type: none"> Basic Skills Off-the-job training Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Portugal <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=35,390 (Intervention: n=10,879; Comparison: n=24,511)</p>

¹⁸ Treatment and comparison group sizes estimated — see Table 9

¹⁹ Treatment and comparison group sizes estimated — see Table 9

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #8: Caliendo et al. (2011)	Preparatory Training (PT)	<ul style="list-style-type: none"> • Basic Skills • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Germany <i>Population with additional barriers:</i> No <i>Study confidence:</i> Medium <i>Sample size:</i> not reported (Intervention: n=1522; Comparison: not reported)²⁰</p>
Study #9: Caliendo et al. (2011)	Short-Term Training (STT)	<ul style="list-style-type: none"> • Basic Skills 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Germany <i>Population with additional barriers:</i> No <i>Study confidence:</i> Medium <i>Sample size:</i> not reported (Intervention: n=2864; Comparison: not reported)²¹</p>
Study #10: Caliendo et al. (2011)	Further Training Measures (FTM)	<ul style="list-style-type: none"> • On-the-job training 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Germany <i>Population with additional barriers:</i> No <i>Study confidence:</i> Medium <i>Sample size:</i> not reported (Intervention: n=924; Comparison: not reported)²²</p>

²⁰ Total and comparison group sample sizes are estimated — see Table 9

²¹ Total and comparison group sample sizes are estimated — see Table 9

²² Total and comparison group sample sizes are estimated — see Table 9

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #11: Cave et al (1993)	JOBSTART Demonstration	<ul style="list-style-type: none"> • Basic Skills • Life Skills • Off-the-job training • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • Hours worked 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> High <i>Sample size:</i> n=1941 (Intervention: n=988; Comparison: n=953)</p>
Study #12: Courtney et al. (2011)	Massachusetts Adolescent Outreach Program	<ul style="list-style-type: none"> • Life Skills • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Other 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Low <i>Sample size:</i> n=179 (Intervention: n=88; Comparison: n=91)</p>
Study #13: Courtney et al. (2019)	YVLifeSet	<ul style="list-style-type: none"> • Life Skills • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> High <i>Sample size:</i> n=1114 (Intervention: n=659; Comparison: n=455)</p>
Study #14: Canzian et al. (2020)	Work experience for young people (WIJ!)	<ul style="list-style-type: none"> • Life Skills • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Belgium <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=68,046 (Intervention: n=4935; Comparison: n=63,111)</p>

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #15: Davis & Heller (2017)	One Summer Chicago Plus — 2012	<ul style="list-style-type: none"> • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings 	<p><i>Design: Randomised</i> <i>Location: United States</i> <i>Population with additional barriers: Yes</i> <i>Study confidence: Medium</i> <i>Sample size: n=1334 (Intervention: n=591; Comparison: n=743)</i></p>
Study #16: Davis & Heller (2017)	One Summer Chicago Plus — 2013	<ul style="list-style-type: none"> • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings 	<p><i>Design: Randomised</i> <i>Location: United States</i> <i>Population with additional barriers: Yes</i> <i>Study confidence: Medium</i> <i>Sample size: n=3742 (Intervention: n=1870; Comparison: n=1872)</i></p>
Study #17: Donato et al. (2018)	Vocational Training, Piedmont	<ul style="list-style-type: none"> • Off-the-job training 	<ul style="list-style-type: none"> • Other 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design: Non-randomised</i> <i>Location: Italy</i> <i>Population with additional barriers: No</i> <i>Study confidence: Low</i> <i>Sample size: n=1217 (Intervention: n=601; Comparison: n=606)</i></p>
Study #18: De Giorgi (2005)	New Deal for Young People	<ul style="list-style-type: none"> • Basic Skills • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design: Non-randomised</i> <i>Location: United Kingdom</i> <i>Population with additional barriers: No</i> <i>Study confidence: Low</i> <i>Sample size: not reported (Intervention: n=895; Comparison: not reported)²³</i></p>

²³ Total and comparison group sample sizes are estimated — see Table 9

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #19: Duarte et al (2020)	Youth Employment Initiative	<ul style="list-style-type: none"> • Off-the-job training • On-the-job training • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings 	<p><i>Design:</i> Non-randomised <i>Location:</i> Portugal <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> not reported (Intervention: n=42,044; Comparison: not reported)²⁴</p>
Study #20: Ehlert et al. (2012a)	Temporary Work ALMP	<ul style="list-style-type: none"> • Basic Skills • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Germany <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=314 (Intervention: n=211; Comparison: n=103)</p>
Study #21: Fein & Hamadyk (2018)	Year Up, Multi-site	<ul style="list-style-type: none"> • Basic Skills • Life Skills • Off-the-job training • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Wages or earnings • Hours worked 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=2496 (Intervention: n=1638; Comparison: n=858)</p>
Study #22: Fraker et al. (2018)	Youth Transition Demonstration Evaluation, Transition WORKS, Erie County, NY	<ul style="list-style-type: none"> • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • Hours worked • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Medium <i>Sample size:</i> n=718 (Intervention: n=397; Comparison: n=321)</p>

²⁴ Total and comparison group sample sizes are estimated — see Table 9

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #23: Fraker et al. (2018)	Youth Transition Demonstration Evaluation, Broadened Horizons, Brighter Futures, Miami-Dade County, NY	<ul style="list-style-type: none"> • Life Skills • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • Hours worked • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Medium <i>Sample size:</i> n=685 (Intervention: n=375; Comparison: n=310)</p>
Study #24: Fraker et al. (2018)	Youth Transition Demonstration Evaluation, YTD, Bronx NY	<ul style="list-style-type: none"> • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • Hours worked • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> High <i>Sample size:</i> n=740 (Intervention: n=420; Comparison: n=320)</p>
Study #25: Fraker et al. (2018)	Youth Transition Demonstration Evaluation, Career Transition Program, Montgomery County, MD	<ul style="list-style-type: none"> • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • Hours worked • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Medium <i>Sample size:</i> n=595 (Intervention: n=320; Comparison: n=275)</p>
Study #26: Fraker et al. (2018)	Youth Transition Demonstration Evaluation, Youth Works, West Virginia	<ul style="list-style-type: none"> • Life Skills • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • Hours worked • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> High <i>Sample size:</i> n=676 (Intervention: n=365; Comparison: n=311)</p>

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #27: Geckeler et al. (2017)	Los Angeles Reconnections Career Academy (LARCA)	<ul style="list-style-type: none"> Life Skills Off-the-job training Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> High <i>Sample size:</i> n=1247 (Intervention: n=649; Comparison: n=598)</p>
Study #28: Gupta et al. (2016)	Linking Innovation, Knowledge, and Employment Program (@LIKE)	<ul style="list-style-type: none"> Basic Skills Life Skills Coaching & Mentoring Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status High school (or equiv.) completion 	<p><i>Design:</i> Non-randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=7387 (Intervention: n=644; Comparison: n=6743)</p>
Study #29: Hämäläinen & Tuomala (2008)	Labour Market Training	<ul style="list-style-type: none"> Basic Skills Off-the-job training 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Finland <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=32,355 (Intervention: n=17,030; Comparison: n=15,325)</p>
Study #30: Hollenbeck and Huang (2006)	High School Career and Technical Education Programmes, Washington — 2006	<ul style="list-style-type: none"> Off-the-job training 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked 	<p><i>Design:</i> Non-randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=51,076 (Intervention: n=25,538; Comparison: 25,538)</p>
Study #31: Hollenbeck and Huang (2006)	Workforce Investment Act, Youth Program, Washington — 2006	<ul style="list-style-type: none"> Coaching & Mentoring Other 	<ul style="list-style-type: none"> Other 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked 	<p><i>Design:</i> Non-randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=10,769 (Intervention: n=5398; Comparison: n=5398)</p>

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #32: Hollenbeck and Huang (2006)	Workforce Investment Act, Apprenticeship Programs — 2006	<ul style="list-style-type: none"> Apprenticeships 	<ul style="list-style-type: none"> Other 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked 	<p><i>Design</i> Non-randomised: <i>Location</i>: United States <i>Population with additional barriers</i>: No <i>Study confidence</i>: Medium <i>Sample size</i>: 10,608 (Intervention: n=5304; Comparison: 5304)</p>
Study #33: Hollenbeck and Huang (2016)	High School Career and Technical Education Programmes, Washington — 2016	<ul style="list-style-type: none"> Off-the-job training 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked 	<p><i>Design</i>: Non-randomised <i>Location</i>: United States <i>Population with additional barriers</i>: No <i>Study confidence</i>: Low <i>Sample size</i>: n=131,708 (Intervention: n=67,520; Comparison: n=64,188)</p>
Study #34: Hollenbeck and Huang (2016)	Workforce Investment Act, Youth Program, Washington — 2016	<ul style="list-style-type: none"> Coaching & Mentoring Other 	<ul style="list-style-type: none"> Other 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked 	<p><i>Design</i>: Non-randomised <i>Location</i>: United States <i>Population with additional barriers</i>: No <i>Study confidence</i>: Medium <i>Sample size</i>: n=6746 (Intervention: n=3373; Comparison: n=3373)</p>
Study #35: Hollenbeck and Huang (2016)	Workforce Investment Act, Apprenticeship Programs — 2016	<ul style="list-style-type: none"> Apprenticeships 	<ul style="list-style-type: none"> Other 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked 	<p><i>Design</i>: Non-randomised <i>Location</i>: United States <i>Population with additional barriers</i>: No <i>Study confidence</i>: Medium <i>Sample size</i>: n=12,572 (Intervention: n=6286; Comparison: n=6286)</p>
Study #36: Izzo et al. (2000)	Extended Transition Services	<ul style="list-style-type: none"> Life Skills Off-the-job training Other 	<ul style="list-style-type: none"> Off-the-job training 	<ul style="list-style-type: none"> Employment status Wages or earnings 	<p><i>Design</i>: Randomised <i>Location</i>: United States <i>Population with additional barriers</i>: Yes <i>Study confidence</i>: Low <i>Sample size</i>: n=47 (Intervention: n=30; Comparison: n=17)</p>

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #37: Jastrzab et al. (1996)	Youth Conservation and Service Corps	<ul style="list-style-type: none"> Life Skills 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Hours worked 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=2382 (Intervention: not reported; Comparison: not reported)</p>
Study #38: Kim et al. (2019)	Independent Living Services	<ul style="list-style-type: none"> Coaching & Mentoring Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status High school (or equiv.) completion 	<p><i>Design:</i> Non-randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Low <i>Sample size:</i> n=4206 (Intervention: n=2757; Comparison: n=1149)</p>
Study #39: Kopečná (2016)	Youth Guarantee	<ul style="list-style-type: none"> On-the-job training 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings 	<p><i>Design:</i> Non-randomised <i>Location:</i> Czechia <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=1503 (Intervention: n=772; Comparison: n=731)</p>
Study #40: Larsson (2003)	Youth Practice	<ul style="list-style-type: none"> Basic Skills Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings 	<p><i>Design:</i> Non-randomised <i>Location:</i> Sweden <i>Population with additional barriers:</i> No <i>Study confidence:</i> Medium <i>Sample size:</i> n=2810 (Intervention: n=606; Comparison: n=2204)</p>
Study #41: Maibom et al. (2014)	Danish Active Labor Market Policies (ALMPs) for Uneducated Youth	<ul style="list-style-type: none"> Basic Skills Coaching & Mentoring Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status 	<p><i>Design:</i> Randomised <i>Location:</i> Denmark <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=2268 (Intervention: n=1115; Comparison: n=1153)</p>

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #42: Maibom et al. (2014)	Danish Active Labor Market Policies (ALMPs) for Educated Youth	<ul style="list-style-type: none"> Coaching & Mentoring Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status 	<p><i>Design:</i> Randomised <i>Location:</i> Denmark <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=1112 (Intervention: n=568; Comparison: n=544)</p>
Study #43: McClanahan et al. (2004)	Summer Career Exploration Program (SCEP)	<ul style="list-style-type: none"> Life Skills Coaching & Mentoring Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=1574 (Intervention: n=1076; Comparison: n=498)</p>
Study #44: Millenky et al. (2014)	National Guard Youth ChalleNGe	<ul style="list-style-type: none"> Coaching & Mentoring Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> High <i>Sample size:</i> n=1173 (Intervention: n=722; Comparison: n=451)</p>
Study #45: Millenky et al. (2018)	YouthBuild	<ul style="list-style-type: none"> Basic Skills Life Skills Off-the-job training Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings High school (or equiv.) completion Vocational Education commencement University commencement 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> High <i>Sample size:</i> n=3929 (Intervention: n=1794; Comparison: n=937)</p>

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #46: Miller et al. (2005)	Centre for Employment Training Replication, San Jose	<ul style="list-style-type: none"> • Basic Skills • Off-the-job training • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • Hours worked • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> High <i>Sample size:</i> n=1136 (Intervention: n=595; Comparison: n=541)</p>
Study #47: Muñoz-Repiso & Braza (2011)	Training Schools Program	<ul style="list-style-type: none"> • Off-the-job training 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Spain <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=225 (Intervention: n=150; Comparison: n=75)</p>
Study #48: Nadon (2020)	Independent Living, Budgeting and Financial Education and Post-Secondary Education Services	<ul style="list-style-type: none"> • Life Skills 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Low <i>Sample size:</i> n=2374 (Intervention: n=1187; Comparison: n=1187)</p>
Study #49: Nadon (2020)	Independent Living, Post-Secondary Education Services	<ul style="list-style-type: none"> • Basic Skills • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Low <i>Sample size:</i> n=2374 (Intervention: n=1187; Comparison: n=1187)</p>
Study #50: Pastore & Pompili (2019)	PIPOL, Training	<ul style="list-style-type: none"> • Off-the-job training 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status 	<p><i>Design:</i> Non-randomised <i>Location:</i> Italy <i>Population with additional barriers:</i> No <i>Study confidence:</i> Medium <i>Sample size:</i> n=10,964 (Intervention: n=1798; Comparison: n=9166)</p>

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #51: Price et al. (2011)	Youth Corps	<ul style="list-style-type: none"> Life Skills Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings University commencement 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=1349 (Intervention: n=935; Comparison: n=414)</p>
Study #52: Quint et al. (1997)	New Chance	<ul style="list-style-type: none"> Basic Skills Life Skills Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked High school (or equiv.) completion Vocational Education commencement 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> High <i>Sample size:</i> n=2079 (Intervention: n=1401; Comparison: n=678)</p>
Study #53: Roder & Elliot (2014)	Year Up, Pilot Study	<ul style="list-style-type: none"> Basic Skills Life Skills Off-the-job training Coaching & Mentoring Other 	<ul style="list-style-type: none"> Services as usual 	<ul style="list-style-type: none"> Employment status Wages or earnings Hours worked 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=164 (Intervention: n=120; Comparison: n=44)</p>
Study #54: Rosholm et al. (2019)	Bridging the Gap between Welfare and Education	<ul style="list-style-type: none"> Basic Skills Life Skills Coaching & Mentoring Other 	<ul style="list-style-type: none"> Off-the-job training On-the-job training Other 	<ul style="list-style-type: none"> Employment status High school (or equiv.) completion 	<p><i>Design:</i> Non-randomised <i>Location:</i> Denmark <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> not reported (Intervention: n=2405; Comparison: n=not reported)</p>

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #55: Schaeffer et al. (2014)	Community Restitution Apprenticeship- Focused Training	<ul style="list-style-type: none"> • Basic Skills • Off-the-job training • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Hours worked • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Low <i>Sample size:</i> n=97 (Intervention: n=50; Comparison: n=47)</p>
Study #56: Schochet et al (2008)	Job Corps	<ul style="list-style-type: none"> • Basic Skills • Off-the-job training • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • Hours worked • High school (or equiv.) completion • Vocational Education commencement • University commencement 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> High <i>Sample size:</i> n=11,313 (Intervention: n=6,828; Comparison: n=4,485)</p>
Study #57: Stromback (2010)	Vocational Education and Training	<ul style="list-style-type: none"> • Off-the-job training 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Wages or earnings 	<p><i>Design:</i> Non-randomised <i>Location:</i> Australia <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> not reported (Intervention: not reported; Comparison: not reported)</p>

REFERENCE	INTERVENTION NAME	INTERVENTION COMPONENTS	COMPARISON COMPONENTS	OUTCOMES	STUDY DETAILS
Study #58: Theodos et al. (2017)	Urban Alliance High School Internship Program	<ul style="list-style-type: none"> • Life Skills • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • High school (or equiv.) completion 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=555 (Intervention: not reported; Comparison: not reported)²⁵</p>
Study #59: Wasserman et al. (2019)	Bridges to Pathways	<ul style="list-style-type: none"> • Life Skills • Coaching & Mentoring • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings • Hours worked • High school (or equiv.) completion • Vocational Education commencement 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> No <i>Study confidence:</i> Low <i>Sample size:</i> n=228 (Intervention: n=137; Comparison: n=91)</p>
Study #60: Wehman et al. (2017)	Project SEARCH, plus ASD supports	<ul style="list-style-type: none"> • On-the-job training • Other 	<ul style="list-style-type: none"> • Services as usual 	<ul style="list-style-type: none"> • Employment status • Wages or earnings 	<p><i>Design:</i> Randomised <i>Location:</i> United States <i>Population with additional barriers:</i> Yes <i>Study confidence:</i> Low <i>Sample size:</i> n=49 (Intervention: n=31; Comparison: n=18)</p>

²⁵ Intervention and comparison group sizes estimated — see Table 9

Appendix B Supplementary information about study methodology

Table 9 Details of decisions made during the transformation of effect sizes

PRIMARY REFERENCE	OUTCOME	IDENTIFIED ISSUE	ACTION
Alegre et al. (2015)	Employment status	Results for both starters (ITT) and completers (TOT) are reported	We used the results for starters (ITT)
	Employment status	Only the sample size for the treatment group was reported	We assumed that that 1:1 matching was used and therefore the size of the control group matched the treatment group
	Employment status	Results are only reported graphically. It was possible to extract the mean difference (and confidence interval) from the reported figure.	Authors were not able to provide results, so these were extracted using a plot digitizer tool (<i>PlotDigitizer Online App</i> , n.d.).
	Employment status	The information was insufficient to transform into a common effect size using functions available in the <i>esc</i> R package.	The mean difference and confidence interval were standardised by dividing by their standard deviation (which was derived from the extracted confidence intervals).
Bloom et al (1993)	Employment status / Education completion	Treatment and control group sample sizes are not reported, only overall sample	Treatment and control group sample sizes are estimated from the ratio reported by the authors (treatment: 2/3, control: 1/3)
Bloom et al (1993)	Employment status / Education completion	Treatment and control group sample sizes are not reported, only overall sample	Treatment and control group sample sizes are estimated from the ratio reported by the authors (treatment: 2/3, control: 1/3)
Bloom et al (1993)	Employment status / Education completion	Treatment and control group sample sizes are not reported, only overall sample	Treatment and control group sample sizes are estimated from the ratio reported by the authors (treatment: 2/3, control: 1/3)
Brunetti & Corsini (2017)	Employment status	The authors report four models based on different matching specifications: stratification matching, radius matching, nearest neighbour matching, kernel matching. All of the results are similar.	We selected the model based on kernel matching.
Centeno et al. (2008)	Employment status	The authors report four models using different DiD specifications: unrestricted DiD, restricted DiD, DiD with PSM (kernel) matching, DiD with PSM (spline) matching. All of the results are similar.	We selected a DiD model using PSM (kernel) matching based on author preferences

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

PRIMARY REFERENCE	OUTCOME	IDENTIFIED ISSUE	ACTION
	Employment status	The authors report inconsistent numbers of participants their summary statistics (Table 1) and model results (Table 2).	After receiving no response from the authors to our query, we have used the numbers reported in Table 1
Caliendo et al. (2011)	Employment status	The number of observations in the control group are not reported	Control group sample sizes are estimated from the average ratio (1:20) reported by the authors
	Employment status	Results for different geographies (states of former East and West Germany) are reported separately	These outcomes were combined by adding percentages
Courtney et al. (2019)	Education completion	High school completion and attainment of general education development results were reported separately	These outcomes were combined by adding percentages
De Giorgi (2005)	Employment status	Male and Female results reported separately	Results for Males and Females were combined in a meta-analysis to obtain a pooled effect for this programme
	Employment status	Treatment and Comparison group sizes are not reported	We assumed that that 1:1 matching was used and therefore the size of the control group matched the treatment group
Duarte et al. (2020)	Employment status	The number of observations in the control group are not reported	We assumed that that 1:1 matching was used and therefore the size of the control group matched the treatment group
	Employment status	The impact of different lengths of on-the-job training are reported (6 months, 12 months and 18 months).	Results for 6 months and 12 months were combined in a meta-analysis to obtain a pooled effect for this programme. Results for 18 months were excluded, as these fall outside our eligibility criteria for this component.
Ehlert et al. (2012a)	Employment status	SE or SD not reported	SE derived from reported information (regression coefficient and t-statistic), SD calculated from SE
Gupta et al. (2016)	Employment status / Education completion	It is not reported at what point in time outcomes are measured	We have assumed 6 months

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

PRIMARY REFERENCE	OUTCOME	IDENTIFIED ISSUE	ACTION
Hämäläinen & Tuomala (2008)	Employment status	Sample size and standard errors were not reported	This information was provided by the authors
Hollenbeck & Huang (2006)	Employment status / Education completion	Results for 2001/02 and 2002/03 were reported separately	These outcomes were combined by adding percentages
Hollenbeck & Huang (2016)	Employment status / Education completion	Results for 2010/11 and 2011/12 were reported separately	These outcomes were combined by adding percentages
Maibom et al (2014)	Employment status	Neither the SE or SD was not reported	SE was derived from reported information (regression coefficient and t-statistic), subsequently the SD was derived from the SE
Nadon (2020)	Employment status	The two studies reported in this study, uses the same data source as Kim (2019)	Kim (2019) was selected as the primary study. Nadon (2020) was excluded from the analysis
Quint et al (1997)	Education completion	High school completion and attainment of general education development results were reported separately	These outcomes were combined by adding percentages
Rosholm et al. (2019)	Employment status / Education completion	Results are only available graphically	Authors were not able to provide results, so these were extracted using a plot digitizer tool (<i>PlotDigitizer Online App</i> , n.d.).
	Employment status / Education completion	Sample size for comparison group is not provided	We assumed that that 1:1 matching was used and therefore the size of the control group matched the treatment group
Theodos et al. (2017)	Employment status	The authors only reported total observations for each analysis	Control and treatment group sizes were estimated from proportions in the treatment and control group

Appendix C Supplementary information about included and excluded studies

Table 10 Distinguishing between multiple reports of the same study

INTERVENTION NAME	PRIMARY REFERENCE	SECONDARY REFERENCE(S)
Inserjovem	Centeno et al. (2008)	Centeno and Novo (2006)
Job Corps	Schochet et al (2008)	Schochet et al (2001) Schochet et al (2006) Zhang et al (2009) Lee et al (2009) Flores-Lagunes (2010) Bampasidou (2012) Bampasidou et al (2014) Frumento et al. (2012) Blanco et al. (2013a) Blanco et al. (2013b) Blanco & Flores-Lagunes (2017) Gritz & Johnson (2001) Chen (2013) Chen et al. (2018)
Job Training Partnership Act (JTPA) — Classroom training	Bloom et al (1993)	Bloom et al. (1997) Heckman et al. (1997) Heckman & Smith (1999) Kornfeld & Bloom (1999)
Job Training Partnership Act (JTPA) — OJT/JSA	Bloom et al (1993)	Bloom et al. (1997) Heckman et al. (1997) Heckman & Smith (1999) Kornfeld & Bloom (1999)
Job Training Partnership Act (JTPA) — Other services	Bloom et al (1993)	Bloom et al. (1997) Heckman et al. (1997) Heckman & Smith (1999) Kornfeld & Bloom (1999)
National Guard Youth ChalleNGe	Millenky et al. (2014)	Millenky et al. (2011) Millenky et al. (2010)
Project SEARCH, plus ASD supports	Wehman et al. (2017)	Wehman et al. (2014)

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

INTERVENTION NAME	PRIMARY REFERENCE	SECONDARY REFERENCE(S)
Temporary Work ALMP	Ehlert et al. (2012a)	Ehlert et al. (2012b)
Urban Alliance High School Internship Program	Theodos et al. (2017)	Theodos et al. (2016)
Year Up, Pilot Study	Roder & Elliot (2014)	Roder & Elliot (2011)
YouthBuild	Millenky et al. (2018)	Miller et al. (2016)
Youth Transition Demonstration Evaluation, Transition WORKS, Erie County, NY	Fraker et al. (2018)	Fraker et al. (2014) Fraker et al. (2011)
Youth Transition Demonstration Evaluation, Broadened Horizons, Brighter Futures, Miami-Dade County, NY	Fraker et al. (2018)	Fraker et al. (2018) Fraker et al. (2014) Fraker et al. (2012)
Youth Transition Demonstration Evaluation, YTDP, Bronx NY	Fraker et al. (2018)	Fraker et al. (2014) Fraker et al. (2011)
Youth Transition Demonstration Evaluation, Career Transition Program, Montgomery County, MD	Fraker et al. (2018)	Fraker et al. (2014) Fraker et al. (2012)
Youth Transition Demonstration Evaluation, Youth Works, West Virginia	Fraker et al. (2018)	Fraker et al. (2014) Fraker et al. (2012)
YVLifeSet	Courtney et al. (2019)	Skemer et al. (2016) Valentine et al. (2015)

Table 11 Selection of studies rejected at full-text review

REFERENCE	INTERVENTION NAME	RATIONALE
Blundell et al. (2004)	New Deal for Young People, Job Assistance elements	Wrong Intervention — intervention components fit under “other”
Caliendo et al. (2011)	Job Search (JS)	Wrong Intervention — intervention components fit under “other”
Caliendo et al. (2011)	Job Creation Schemes (JCS)	Wrong Intervention — intervention components fit under “other”

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

REFERENCE	INTERVENTION NAME	RATIONALE
Cumming et al. (2018) Skemer et al. (2017)	Young Adult Internship Program	Wrong Intervention — intervention components fit under “other”
Fraker et al. (2018) Fraker et al. (2014) Fraker et al. (2011)	Youth Transition Demonstration Evaluation, Youth Wins, Colorado	Wrong Intervention — intervention components fit under “other”
Pastore & Pompili (2019)	PIPOL, Internships	Wrong Intervention — intervention components fit under “other”
Rotar (2012b) Rotar (2012a)	Slovenian Institutional Training Program	Wrong Intervention — intervention components fit under “other”
Zinn & Courtney (2017) Courtney et al. (2011)	Independent Living, Employment Services, Kern County CA	Wrong Intervention — intervention components fit under “other”

Table 12 Details of included studies reporting hours worked

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
Bloom et al (1993)	Job Training Partnership Act (JTPA) — Classroom training	Hours worked at Quarter 6 (Female, Male Youth)	SD or SE not reported
Bloom et al (1993)	Job Training Partnership Act (JTPA) — OJT/JSA	Hours worked at Quarter 6 (Female, Male Youth)	SD or SE not reported
Bloom et al (1993)	Job Training Partnership Act (JTPA) — Other services	Hours worked at Quarter 6 (Female, Male Youth)	SD or SE not reported
Cave et al. (1993)	JOBSTART Demonstration	Total Hours worked Year 4	SD or SE not reported
Fein & Hamadyk (2018)	Year Up Multi-Site	Average weekly hours worked at time of 18-month follow-up survey	SE reported, SD derived
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — Transition WORKS (Erie, NY)	Total hours worked in paid job in the last year (36-month survey)	SD or SE not reported

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — Broadened Horizons, Brighter Futures	Total hours worked in paid job in the last year (36-month survey)	SD or SE not reported
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — YTDP	Total hours worked in paid job in the last year (36-month survey)	SD or SE not reported
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — Career Transition Program	Total hours worked in paid job in the last year (36-month survey)	SD or SE not reported
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — Youth Works	Total hours worked in paid job in the last year (36-month survey)	SD or SE not reported
Hollenbeck & Huang (2006)	High school career and technical education programmes	Average quarterly hours 3 quarters following program exit	SD or SE not reported
Hollenbeck & Huang (2006)	Workforce Investment Act (Youth Program)	Average quarterly hours 3 quarters following program exit	SD or SE not reported
Hollenbeck & Huang (2006)	Workforce Investment Act — Apprenticeship Programs	Average quarterly hours 3 quarters following program exit	SD or SE not reported
Hollenbeck & Huang (2016)	High school career and technical education programmes	Average quarterly hours 3 quarters following program exit	SD or SE not reported
Hollenbeck & Huang (2016)	Workforce Investment Act (Youth Program)	Average quarterly hours 3 quarters following program exit	SD or SE not reported
Hollenbeck & Huang (2016)	Workforce Investment Act — Apprenticeship Programs	Average quarterly hours 3 quarters following program exit	SD or SE not reported
McClanahan et al. (2004)	SCEP	Average hours worked for three-month period	SD or SE not reported
Miller et al. (2005)	Centre for Employment Training	Number of months worked Year 1 / 2 / 3 / 4	SD or SE not reported

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
Quint et al (1997)	New Chance	Average hours worked 31-42 months follow-up	SD or SE not reported
Roder & Elliot (2014)	Year Up Pilot	Number of hours worked during the 4th year after random assignment	SD or SE not reported
Schaeffer et al. (2014)	Community Restitution Apprenticeship-Focused Training	Hours worked per month	SD reported
Schochet et al (2008)	Job Corps	Average hours employed per week in Year 4	SD or SE not reported
Wasserman et al. (2019)	Bridges to Pathways Program	Hours worked per week (among those who were employed)	SD or SE not reported

Table 13 Details of included studies reporting wages or earnings

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
Bauer et al. (2014)	New York City Justice Corps	Average cumulative wages after 24 months	SD or SE not reported
Bloom et al (1993)	Job Training Partnership Act (JTPA) — Classroom training	Earnings at quarter 6, Earnings over 18-month period	SD or SE not reported
Bloom et al (1993)	Job Training Partnership Act (JTPA) — OJT/JSA	Earnings at quarter 6, Earnings over 18-month period	SD or SE not reported
Bloom et al (1993)	Job Training Partnership Act (JTPA) — Other services	Earnings at quarter 6, Earnings over 18-month period	SD or SE not reported
Cave et al. (1993)	JOBSTART Demonstration	Total earnings Years 4	SD or SE not reported
Courtney et al. (2011)	Massachusetts Adolescent Outreach Programme	Earnings in the 12 months prior (~2 years after commencement)	ES can be transformed
Courtney et al. (2019)	YVLifeSet	Earnings from formal work (in the year after intervention start)	ES reported

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
Davis & Heller (2017)	One Summer Chicago Plus (OSC+): First RCT, 2016	Earnings two years after program	ES can be transformed
Davis & Heller (2017)	One Summer Chicago Plus (OSC+): Second RCT, 2018	Earnings two years after program	ES can be transformed
Duarte et al. (2020)	Youth Employment Initiative	Effect on wage in 36 months	ES can be transformed
Fein & Hamadyk (2018)	Year Up Multi-Site	Total earnings in Year 3 post-random assignment	ES can be transformed
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — Transition WORKS (Erie, NY)	Total earnings in the past year (36-month survey)	SD or SE not reported
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — Broadened Horizons, Brighter Futures	Total earnings in the past year (36-month survey)	SD or SE not reported
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — YTD-P	Total earnings in the past year (36-month survey)	SD or SE not reported
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — Career Transition Program	Total earnings in the past year (36-month survey)	SD or SE not reported
Fraker et al. (2018)	Youth Transition Demonstration (YTD) — Youth Works	Total earnings in the past year (36-month survey)	SD or SE not reported
Geckeler et al. (2017)	Los Angeles Reconnections Career Academy (LARCA)	Total earnings 2 years since random assignment	SD or SE not reported
Hämäläinen & Tuomala (2008)	Labour Market Training	Earnings two years after program starts (SEK)	ES can be transformed
Hollenbeck & Huang (2016)	High school career and technical	Average quarterly earnings 3 quarters following program exit	SD or SE not reported

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
	education programmes		
Hollenbeck & Huang (2016)	Workforce Investment Act (Youth Program)	Average quarterly earnings 3 quarters following program exit	SD or SE not reported
Hollenbeck & Huang (2016)	Workforce Investment Act — Apprenticeship Programs	Average quarterly earnings 3 quarters following program exit	SD or SE not reported
Hollenbeck & Huang (2006)	High school career and technical education programmes	Average quarterly earnings 3 quarters following program exit	SD or SE not reported
Hollenbeck & Huang (2006)	Workforce Investment Act (Youth Program)	Average quarterly earnings 3 quarters following program exit	SD or SE not reported
Hollenbeck & Huang (2006)	Workforce Investment Act — Apprenticeship Programs	Average quarterly earnings 3 quarters following program exit	SD or SE not reported
Izzo et al. (2000)	Extended Transition Services	Mean earnings 8 quarters (2 years) following exit from program	ES can be transformed
Kopečná (2016)	Youth Guarantee	Difference in monthly income 18 months post intervention start	ES can be transformed
McClanahan et al. (2004)	SCEP	Average earnings for three-month period	SD or SE not reported
Millenky et al. (2014)	National Guard Youth Challenge	Earnings in last 12 months (36 months following program start)	ES reported
Millenky et al. (2018)	YouthBuild	Earnings in year four since randomisation	SD or SE not reported
Miller et al. (2005)	Centre for Employment Training	Total earnings during 54 month follow up (\$)	SD or SE not reported
Price et al. (2011)	Youth Corps	Total personal income in last year (18-month follow up)	SD or SE not reported
Quint et al (1997)	New Chance	Average total earnings 31-42 months follow-up	SD or SE not reported

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
Roder & Elliot (2014)	Year Up Pilot	Earnings during the 4th year after random assignment	SD or SE not reported
Schochet et al (2008)	Job Corps	Average earnings per week by Year 4	SD or SE not reported
Stromback (2010)	Vocational education and training	Log weekly earnings (full-time only), wave 10 (approx. 23 years old)	SD, SE, or sample size not reported
Theodos et al. (2017)	Urban Alliance Program	Post program wages (24-month survey)	ES can be transformed
Wasserman et al. (2019)	Bridges to Pathways Program	Hourly wage (mean)	SD or SE not reported
Wehman et al. (2017)	Project SEARCH Plus ASD Support	Wages (change from baseline to 12 months post-graduation)	ES can be transformed

Table 14 Details of included studies reporting vocational education commencement

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
Bauer et al. (2014)	New York City Justice Corps	Obtained technical training certificate or license (at 12 month follow up)	Not attempted
Millenky et al. (2014)	Youth Build	Received a trade licence or training certificate within 12 months after programme start	Not attempted
Quint et al (1997)	New Chance	Received trade license by end of month 6 / 18 / 30 / 42	Not attempted
Schochet et al. (2008)	Job Corps	Attained vocational, technical or trade certificate during 48-month period	Not attempted
Wasserman et al. (2019)	Bridges to Pathways Program	Earned professional license or certification (within 12 months)	Not attempted

Table 15 Details of included studies reporting university commencement

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
Schochet et al. (2008)	Job Corps	Attained college degree during 48 month period	Not attempted

A network meta-analysis of employment and skills programmes and interventions designed to assist young people to enter the labour market in high income countries

REFERENCE	INTERVENTION NAME	OUTCOME	ES TRANSFORMATION
Millenky et al. (2014)	Youth Build	Received a post-secondary degree / Associate's degree / Bachelor's degree / other degree within 12 months after programme start	Not attempted
Price et al. (2011)	Youth Corps	Associate's degree or above / Bachelor's degree or above / Graduate degree within 30 month	Not attempted

Appendix D Supplementary NMA results

Figure 29 Forest plot depicting results of Interaction, Additive and standard NMA of component combinations on employment status

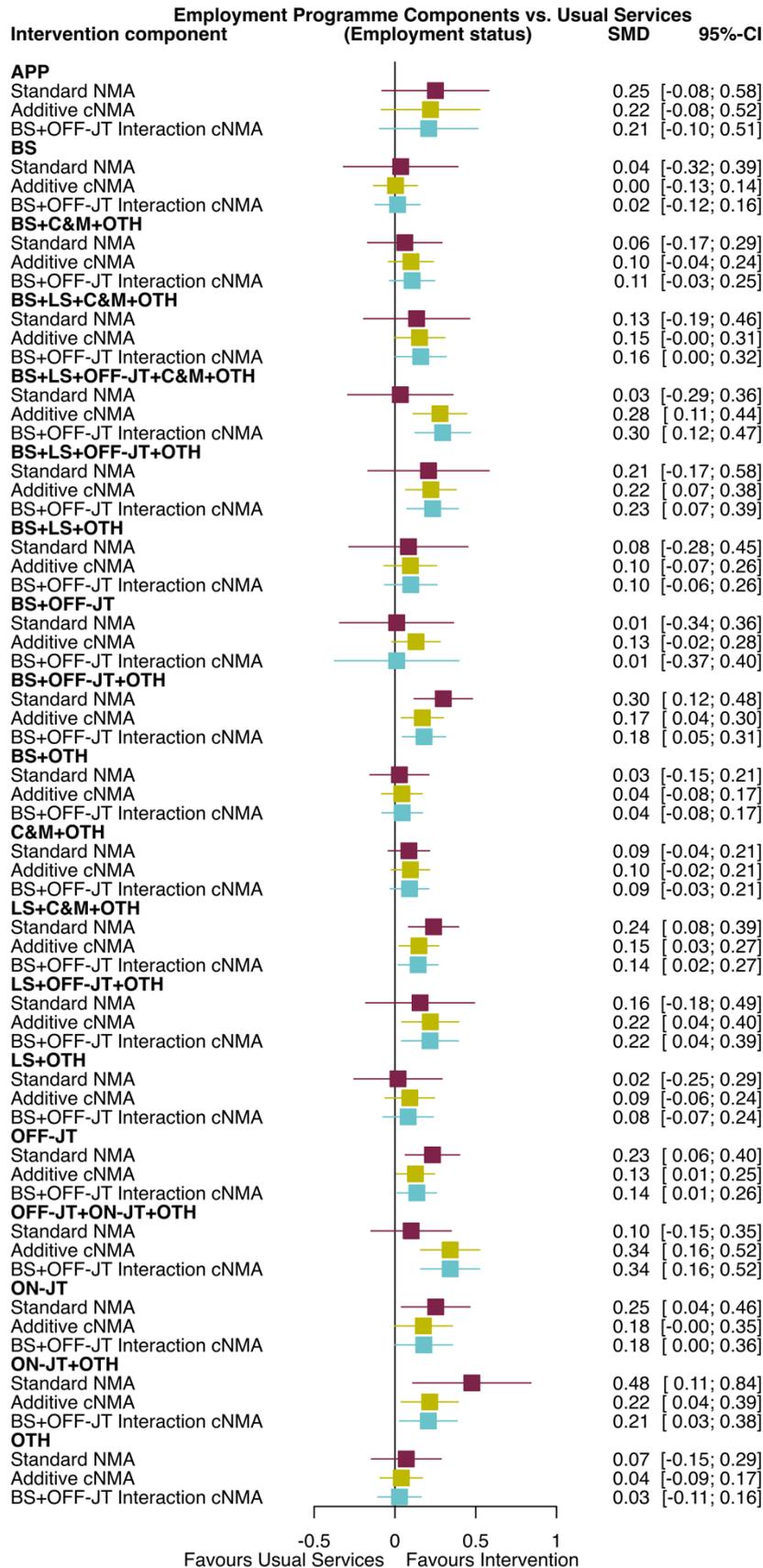


Figure 30 Forest plot depicting results of Interaction, Additive and standard NMA of component combinations on education completion

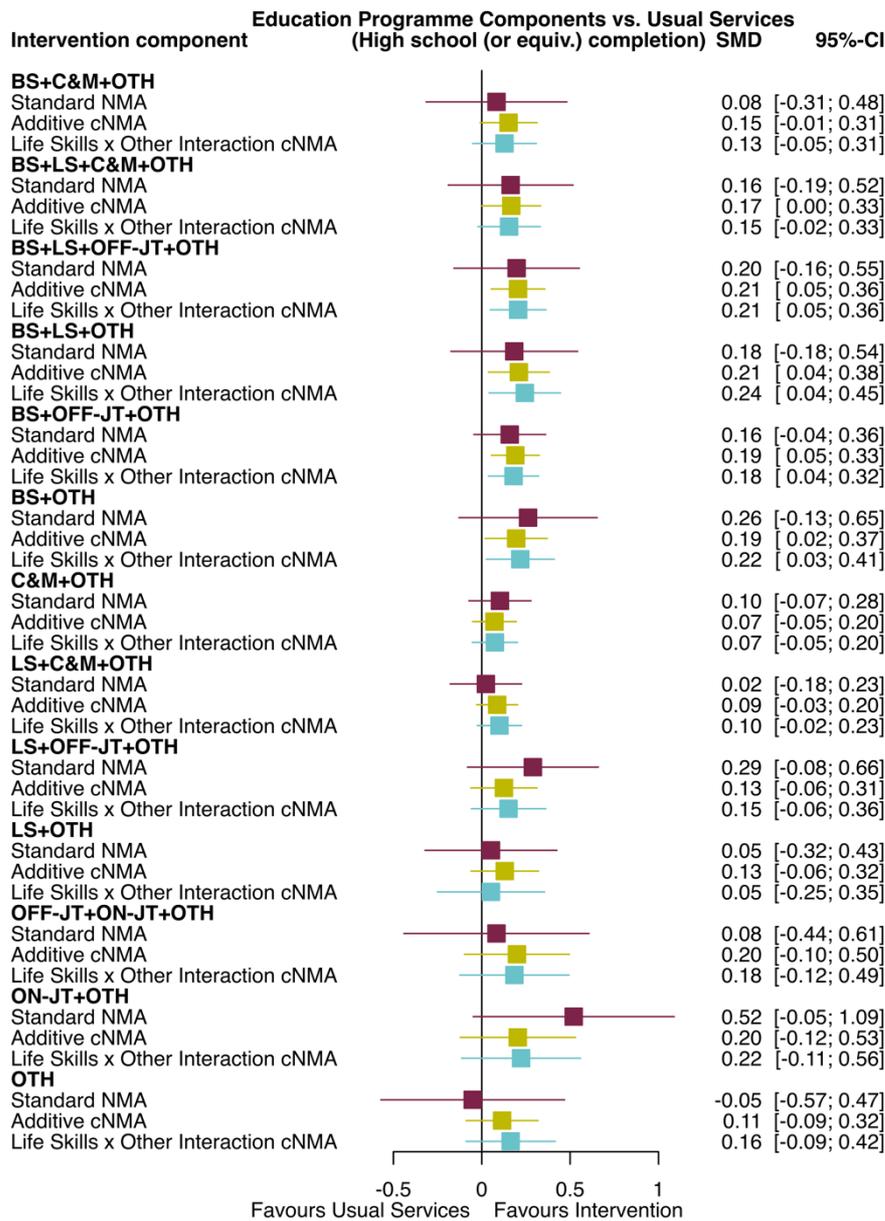


Table 16 Relative effects of combinations of components of employment and skills programmes on employment status²⁶

Direct evidence from pairwise comparisons																				
APP	0.18 [-0.07; 0.43]	.	
0.22 [-0.27; 0.70]	BS	0.04 [-0.32; 0.39]	
0.19 [-0.21; 0.59]	-0.03 [-0.45; 0.40]	BS+C&M+OTH	0.06 [-0.17; 0.29]	
0.12 [-0.35; 0.58]	-0.10 [-0.58; 0.38]	-0.07 [-0.47; 0.33]	BS+LS+C&M+OTH	0.10 [-0.64; 0.84]	.	0.12 [-0.24; 0.48]	
0.22 [-0.25; 0.68]	0.00 [-0.48; 0.48]	0.03 [-0.37; 0.43]	0.10 [-0.36; 0.56]	BS+LS+OFF-JT+C&M+OTH	0.03 [-0.29; 0.36]	
0.04 [-0.46; 0.54]	-0.17 [-0.69; 0.34]	-0.15 [-0.59; 0.29]	-0.07 [-0.57; 0.42]	-0.17 [-0.67; 0.32]	BS+LS+OFF-JT+OTH	0.21 [-0.17; 0.58]	
0.17 [-0.33; 0.66]	-0.05 [-0.56; 0.46]	-0.02 [-0.46; 0.41]	0.05 [-0.44; 0.54]	-0.05 [-0.54; 0.44]	0.12 [-0.40; 0.65]	BS+LS+OTH	0.08 [-0.28; 0.45]	
0.24 [-0.24; 0.72]	0.02 [-0.47; 0.52]	0.05 [-0.37; 0.47]	0.12 [-0.36; 0.61]	0.02 [-0.46; 0.50]	0.20 [-0.32; 0.71]	0.07 [-0.44; 0.58]	BS+OFF-JT	0.01 [-0.34; 0.36]	
-0.05 [-0.43; 0.33]	-0.26 [-0.66; 0.13]	-0.24 [-0.53; 0.05]	-0.17 [-0.54; 0.21]	-0.27 [-0.64; 0.11]	-0.09 [-0.51; 0.32]	-0.22 [-0.63; 0.19]	-0.29 [-0.68; 0.11]	BS+OFF-JT+OTH	0.30 [0.12; 0.48]	
0.22 [-0.15; 0.60]	0.01 [-0.39; 0.41]	0.03 [-0.26; 0.33]	0.11 [-0.27; 0.48]	0.01 [-0.37; 0.38]	0.18 [-0.24; 0.60]	0.06 [-0.35; 0.47]	-0.02 [-0.41; 0.38]	0.27 [0.02; 0.53]	BS+OTH	0.03 [-0.15; 0.21]	
0.16 [-0.16; 0.49]	-0.05 [-0.43; 0.33]	-0.03 [-0.29; 0.24]	0.05 [-0.30; 0.40]	-0.05 [-0.40; 0.30]	0.12 [-0.27; 0.52]	-0.00 [-0.39; 0.39]	-0.08 [-0.45; 0.30]	0.21 [-0.01; 0.43]	-0.06 [-0.28; 0.16]	C&M+OTH	0.02 [-0.23; 0.27]	0.08 [-0.05; 0.22]
0.01 [-0.34; 0.37]	-0.20 [-0.59; 0.18]	-0.18 [-0.46; 0.10]	-0.11 [-0.47; 0.26]	-0.21 [-0.57; 0.15]	-0.03 [-0.44; 0.37]	-0.16 [-0.55; 0.24]	-0.23 [-0.61; 0.16]	0.06 [-0.18; 0.30]	-0.21 [-0.45; 0.03]	-0.15 [-0.35; 0.04]	LS+C&M+OTH	-0.02 [-0.56; 0.51]	0.26 [0.10; 0.42]
0.10 [-0.37; 0.56]	-0.12 [-0.61; 0.37]	-0.09 [-0.50; 0.31]	-0.02 [-0.49; 0.45]	-0.12 [-0.59; 0.35]	0.05 [-0.45; 0.56]	-0.07 [-0.57; 0.43]	-0.14 [-0.63; 0.34]	0.14 [-0.24; 0.53]	-0.13 [-0.51; 0.25]	-0.07 [-0.43; 0.29]	0.08 [-0.29; 0.45]	LS+OFF-JT+OTH	0.80 [0.09; 1.51]	-0.09 [-0.47; 0.29]
0.23 [-0.20; 0.66]	0.02 [-0.43; 0.46]	0.04 [-0.31; 0.40]	0.12 [-0.31; 0.54]	0.01 [-0.41; 0.44]	0.19 [-0.27; 0.65]	0.06 [-0.39; 0.52]	-0.01 [-0.45; 0.44]	0.28 [-0.05; 0.61]	0.01 [-0.32; 0.34]	0.07 [-0.23; 0.37]	0.22 [-0.09; 0.53]	0.14 [-0.30; 0.57]	LS+OTH	0.02 [-0.25; 0.29]
0.02 [-0.33; 0.36]	-0.20 [-0.59; 0.19]	-0.17 [-0.46; 0.11]	-0.10 [-0.47; 0.27]	-0.20 [-0.57; 0.17]	-0.02 [-0.44; 0.39]	-0.15 [-0.55; 0.26]	-0.22 [-0.61; 0.17]	0.07 [-0.18; 0.31]	-0.20 [-0.45; 0.04]	-0.15 [-0.35; 0.06]	0.01 [-0.22; 0.23]	-0.08 [-0.44; 0.28]	-0.21 [-0.53; 0.11]	OFF-JT	0.24 [-0.13; 0.61]	0.27 [0.09; 0.46]
0.15 [-0.26; 0.56]	-0.06 [-0.50; 0.37]	-0.04 [-0.38; 0.30]	0.03 [-0.35; 0.42]	-0.07 [-0.48; 0.34]	0.11 [-0.34; 0.56]	-0.02 [-0.46; 0.43]	-0.09 [-0.52; 0.34]	0.20 [-0.11; 0.51]	-0.07 [-0.38; 0.24]	-0.01 [-0.29; 0.27]	0.14 [-0.15; 0.43]	0.06 [-0.36; 0.47]	-0.08 [-0.45; 0.29]	0.13 [-0.17; 0.43]	OFF-JT+ON-JT+OTH	0.11 [-0.15; 0.37]
-0.00 [-0.39; 0.39]	-0.22 [-0.63; 0.20]	-0.19 [-0.50; 0.12]	-0.12 [-0.51; 0.27]	-0.22 [-0.61; 0.17]	-0.04 [-0.48; 0.39]	-0.17 [-0.59; 0.26]	-0.24 [-0.65; 0.17]	0.05 [-0.23; 0.33]	-0.22 [-0.50; 0.05]	-0.17 [-0.41; 0.08]	-0.01 [-0.28; 0.25]	-0.10 [-0.49; 0.30]	-0.23 [-0.58; 0.11]	-0.02 [-0.29; 0.25]	-0.15 [-0.48; 0.17]	ON-JT	.	.	.	0.25 [0.04; 0.46]
-0.23 [-0.72; 0.27]	-0.44 [-0.95; 0.07]	-0.42 [-0.85; 0.02]	-0.34 [-0.83; 0.15]	-0.44 [-0.93; 0.05]	-0.27 [-0.79; 0.26]	-0.39 [-0.91; 0.13]	-0.47 [-0.97; 0.04]	-0.18 [-0.58; 0.23]	-0.45 [-0.86; 0.04]	-0.39 [-0.78; 0.00]	-0.24 [-0.63; 0.16]	-0.32 [-0.82; 0.18]	-0.46 [-0.91; 0.00]	-0.24 [-0.65; 0.16]	-0.38 [-0.82; 0.07]	-0.22 [-0.65; 0.20]	ON-JT+OTH	.	.	0.48 [0.11; 0.84]
0.18 [-0.07; 0.43]	-0.03 [-0.45; 0.38]	-0.01 [-0.32; 0.31]	0.06 [-0.33; 0.46]	-0.04 [-0.43; 0.35]	0.14 [-0.29; 0.57]	0.01 [-0.41; 0.44]	-0.06 [-0.47; 0.35]	0.23 [-0.05; 0.51]	-0.04 [-0.33; 0.24]	0.02 [-0.19; 0.22]	0.17 [-0.08; 0.42]	0.09 [-0.31; 0.48]	-0.05 [-0.40; 0.30]	0.16 [-0.08; 0.40]	0.03 [-0.30; 0.36]	0.18 [-0.12; 0.49]	0.41 [-0.02; 0.83]	OTH	.	
0.25 [-0.08; 0.58]	0.04 [-0.32; 0.39]	0.06 [-0.17; 0.29]	0.13 [-0.19; 0.46]	0.03 [-0.29; 0.36]	0.21 [-0.17; 0.58]	0.08 [-0.28; 0.45]	0.01 [-0.34; 0.36]	0.30 [0.12; 0.48]	0.03 [-0.15; 0.21]	0.09 [-0.04; 0.21]	0.24 [0.08; 0.39]	0.16 [-0.18; 0.49]	0.02 [-0.25; 0.29]	0.23 [0.06; 0.40]	0.10 [-0.15; 0.35]	0.25 [0.04; 0.46]	0.48 [0.11; 0.84]	0.07 [-0.15; 0.29]	SAU	
Indirect evidence from the network meta-analysis																				

²⁶ Effect sizes (Hedge's g) are reported with 95% confidence intervals. Combinations of components (listed in alphabetical order) from included studies are shown in purple. Direct evidence (i.e., sourced from pairwise comparisons) are shown in aqua. Indirect effects (i.e., from the NMA) are shown in green. Results that are statistically significant at the 95 per cent level are presented in bold. **Plot legend** — BS: Basic Skills, LS: Life Skills, OFF-JT: Off-the-job training, ON-JT: On-the-job-training, APP: Apprenticeships, C&M: Coaching and mentoring, OTH: Other (residual) component.

Table 17 Relative effects of combinations of components of employment and skills programmes on education completion²⁷

Direct evidence from pairwise comparisons													
BS+C&M+OTH	0.08 [-0.31; 0.48]
-0.08 [-0.61; 0.45]	BS+LS+C&M+OTH	0.08 [-0.31; 0.47]	.	.	0.16 [-0.19; 0.52]
-0.11 [-0.65; 0.42]	-0.03 [-0.53; 0.47]	BS+LS+OFF-JT+OTH	0.20 [-0.16; 0.55]
-0.10 [-0.64; 0.44]	-0.02 [-0.52; 0.48]	0.01 [-0.49; 0.52]	BS+LS+OTH	0.18 [-0.18; 0.54]
-0.07 [-0.52; 0.37]	0.01 [-0.40; 0.41]	0.04 [-0.37; 0.45]	0.03 [-0.39; 0.44]	BS+OFF-JT+OTH	0.16 [-0.04; 0.36]
-0.18 [-0.74; 0.38]	-0.10 [-0.63; 0.43]	-0.06 [-0.59; 0.46]	-0.08 [-0.61; 0.45]	-0.10 [-0.54; 0.34]	BS+OTH	0.26 [-0.13; 0.65]
-0.02 [-0.45; 0.42]	0.06 [-0.33; 0.46]	0.09 [-0.30; 0.49]	0.08 [-0.32; 0.48]	0.06 [-0.21; 0.32]	0.16 [-0.27; 0.59]	C&M+OTH	0.10 [-0.07; 0.28]
0.06 [-0.39; 0.51]	0.14 [-0.27; 0.55]	0.17 [-0.23; 0.58]	0.16 [-0.25; 0.57]	0.13 [-0.15; 0.42]	0.24 [-0.20; 0.68]	0.08 [-0.19; 0.35]	LS+C&M+OTH	0.07 [-0.40; 0.55]	0.02 [-0.18; 0.23]
-0.21 [-0.75; 0.34]	-0.13 [-0.64; 0.39]	-0.09 [-0.61; 0.42]	-0.11 [-0.62; 0.41]	-0.13 [-0.55; 0.29]	-0.03 [-0.57; 0.51]	-0.19 [-0.60; 0.22]	-0.27 [-0.69; 0.16]	LS+OFF-JT+OTH	0.29 [-0.08; 0.66]
0.03 [-0.51; 0.58]	0.11 [-0.40; 0.63]	0.15 [-0.37; 0.66]	0.13 [-0.39; 0.65]	0.11 [-0.32; 0.53]	0.21 [-0.33; 0.75]	0.05 [-0.36; 0.46]	-0.03 [-0.45; 0.40]	0.24 [-0.29; 0.76]	LS+OTH	.	.	.	0.05 [-0.32; 0.43]
0.00 [-0.66; 0.66]	0.08 [-0.31; 0.47]	0.11 [-0.52; 0.75]	0.10 [-0.53; 0.74]	0.07 [-0.49; 0.64]	0.18 [-0.48; 0.83]	0.02 [-0.53; 0.57]	-0.06 [-0.62; 0.50]	0.21 [-0.44; 0.85]	-0.03 [-0.67; 0.61]	OFF-JT+ON-JT+OTH	.	.	.
-0.44 [-1.13; 0.26]	-0.36 [-1.03; 0.31]	-0.32 [-0.99; 0.35]	-0.34 [-1.01; 0.34]	-0.36 [-0.97; 0.24]	-0.26 [-0.95; 0.43]	-0.42 [-1.01; 0.18]	-0.50 [-1.10; 0.11]	-0.23 [-0.91; 0.45]	-0.47 [-1.15; 0.21]	-0.44 [-1.21; 0.34]	ON-JT+OTH	.	0.52 [-0.05; 1.09]
0.13 [-0.52; 0.79]	0.21 [-0.41; 0.84]	0.25 [-0.38; 0.88]	0.24 [-0.40; 0.87]	0.21 [-0.35; 0.77]	0.31 [-0.34; 0.96]	0.15 [-0.39; 0.70]	0.07 [-0.40; 0.55]	0.34 [-0.30; 0.98]	0.10 [-0.54; 0.74]	0.13 [-0.60; 0.87]	0.57 [-0.20; 1.34]	OTH	.
0.08 [-0.31; 0.48]	0.16 [-0.19; 0.52]	0.20 [-0.16; 0.55]	0.18 [-0.18; 0.54]	0.16 [-0.04; 0.36]	0.26 [-0.13; 0.65]	0.10 [-0.07; 0.28]	0.02 [-0.18; 0.23]	0.29 [-0.08; 0.66]	0.05 [-0.32; 0.43]	0.08 [-0.44; 0.61]	0.52 [-0.05; 1.09]	-0.05 [-0.57; 0.47]	SAU
Indirect evidence from the network meta-analysis													

²⁷ Effect sizes (Hedge's g) are reported with 95% confidence intervals. Combinations of components (listed in alphabetical order) from included studies are shown in purple. Direct evidence (i.e., sourced from pairwise comparisons) are shown in aqua. Indirect effects (i.e., from the NMA) are shown in green. Results that a statistically significant at the 95 per cent level are presented in bold. **Plot legend** — BS: Basic Skills, LS: Life Skills, OFF-JT: Off-the-job training, ON-JT: On-the-job-training, APP: Apprenticeships, C&M: Coaching and mentoring, OTH: Other (residual) component.

Appendix E Assessing Network Coherence

Evaluating local incoherence

Employment status

Application of the SIDE method to the review team's preferred specification for employment status — visualised in Figure 31 — allows for the exploration of inconsistency within each combination of components where mixed evidence is present.

The analysis established that there are 190 possible combinations of components in this specification. Of those combinations, twelve (n=12) solely use direct evidence. Another eleven (n=11) use mixed evidence i.e., a combination of direct and indirect evidence.

Note that the number of studies for each combination of components using mixed evidence is very small — there are only five combinations of components that use two or more studies — as a result, care needs to be taken into drawing conclusions from any observed inconsistency.

Application of the SIDE method identified a moderate amount of overall heterogeneity ($\tau = 0.180$). Inconsistency between indirect and direct evidence was assessed using I^2 . Inconsistency appears to be present in three combinations, as indicated by an I^2 value of 80 or more:

- Life Skills + Coaching & Mentoring + Other versus Services as Usual (I^2 : 98%)
- Off-the-job Training versus Services as Usual (I^2 97%)
- Off-the-job Training + On-the-job Training + Other versus Services as Usual (I^2 : 88%)

In interpreting these results its important to note that I^2 is a relative — as opposed to absolute — measure of heterogeneity (Borenstein et al., 2017). Therefore taken together, these results suggest that of the moderate amount of heterogeneity that was identified, most it is true heterogeneity.

Education completion

It was not possible to apply the SIDE methodology to the network of studies reporting education completion outcomes due to the absence of any indirect evidence within that network.

Evaluating global incoherence

Employment status

The 'between designs' decomposition of Cochran's Q for employment status is presented in Table 18 below. The results suggest that the level of heterogeneity within the included designs (i.e., combinations of components) is not significant. However, caution should be taken in the interpretation of these results as the number of included studies with mixed evidence is low.

Table 18 Cochran's Q for employment status

SPECIFICATION	Q	df	p	Square root of between study variance
Between designs	8.373	4	0.079	0.180

Education completion

It was not possible to assess global incoherence for the network of included studies reporting education outcomes due to the absence of any indirect evidence within that network.

Figure 31 Forest plot visualising inconsistency within the network for employment status

