Evaluation of

Roots of Empathy
in Scotland 2014-15

Final Report

for Action for Children

Qa Research
Mill House,
North Street,
York, YO1 6JD
01904 632039

Dephna House, 24-26 Arcadia Ave,
London, N3 2JU
0208 8191397

www.qaresearch.co.uk
Company registration: 3186539
# Contents

Contents ........................................................................................................................................ 2
Acknowledgements ......................................................................................................................... 4
Executive summary .......................................................................................................................... 5
1. Introduction and aims .................................................................................................................. 14
2. Methodology ............................................................................................................................. 15
   2.1 Designing a robust evaluation methodology .............................................................................. 15
   2.2 Sampling and selection of research participants ......................................................................... 17
       2.2.1 Sampling of local authorities, schools and classes ............................................................... 17
       2.2.2 Sampling for the teacher report questionnaire .................................................................... 17
       2.2.3 Sampling for the pupil interview ......................................................................................... 18
       2.2.4 Selection of stakeholders for in-depth interview ................................................................. 18
       2.2.5 Selection of parents for focus group or interview ............................................................... 19
   2.3 Research tool design ............................................................................................................... 20
       2.3.1 Teacher-report questionnaire (quantitative) ........................................................................ 20
       2.3.2 Pupil interview (quantitative) .............................................................................................. 22
       2.3.3 Stakeholder interview (qualitative) ...................................................................................... 23
       2.3.4 Parent focus group or interview (qualitative) ..................................................................... 23
   2.4 Data collection ....................................................................................................................... 24
       2.4.1 Baseline teacher questionnaire ............................................................................................ 24
       2.4.2 Follow-up teacher questionnaire ........................................................................................ 24
       2.4.3 Baseline pupil interview .................................................................................................... 25
       2.4.4 Follow-up pupil interview ................................................................................................ 25
       2.4.5 Follow-up stakeholder interview ....................................................................................... 25
       2.4.6 Follow-up parent focus group and in-depth interview ......................................................... 26
3. Key findings ................................................................................................................................ 27
   3.1 Quantitative findings from teacher questionnaire ..................................................................... 27
       3.1.1 Achieved sample .................................................................................................................... 28
       3.1.2 Reliability of scores .............................................................................................................. 28
       3.1.3 Changes in the four variables from pre- to post-, by intervention and control groups ...... 29
       3.1.4 Examining pupils in need: the bottom 25% ......................................................................... 37
       3.1.5 Differences by gender .......................................................................................................... 44
       3.1.6 Associations by age ............................................................................................................ 48
       3.1.7 Empathy as a mediator of change in aggression and prosocial behaviour ....................... 51
   3.2 Quantitative findings from administering KEDS protocol with pupils ...................................... 55
       3.2.1 Achieved sample .................................................................................................................... 55
       3.2.2 Changes in the three variables (affective empathy, cognitive empathy, behavioural empathy) from pre- to post-, by intervention and control ........................................................................... 56
   3.3 Qualitative findings from class teachers, stakeholders & parents ............................................ 64
       3.3.1 Profile of participants .......................................................................................................... 64
       3.3.2 Expectations of the programme ........................................................................................... 66
       3.3.3 Impact of the programme on pupils’ empathy ................................................................. 66
       3.3.4 Impact of the programme on pupils’ aggression ............................................................... 67
       3.3.5 Impact of the programme on pupils’ prosocial behaviour .................................................. 68
       3.3.6 Other impact on pupils ......................................................................................................... 69
       3.3.7 Sustained impact on pupils .................................................................................................. 72
       3.3.8 Impact on other people ........................................................................................................ 73
       3.3.9 Attributing impact to Roots of Empathy .............................................................................. 75
       3.3.10 Process evaluation – what makes Roots of Empathy work .............................................. 76
       3.3.11 Process evaluation – suggested changes to programme delivery .................................... 78
3.3.12 Overall assessment of the Roots of Empathy programme ........................................ 81
3.4 Limitations .................................................................................................................. 82
  3.4.1 Sampling .............................................................................................................. 82
  3.4.2 Confounding ‘ripple effect’ of intervention on control pupils .......................... 82
  3.4.3 Teacher-report questionnaire ........................................................................... 83
  3.4.4 Pupil interview ................................................................................................... 84
  3.4.5 Future potential research .................................................................................... 84

4. Conclusions .................................................................................................................. 85
5. Appendices ................................................................................................................... 87
  5.1 Research tools ......................................................................................................... 87
    5.1.1 Baseline teacher report questionnaire .............................................................. 87
    5.1.2 Follow-up teacher report questionnaire .......................................................... 88
    5.1.3 Baseline pupil interview .................................................................................... 90
    5.1.4 Follow-up pupil interview ............................................................................... 97
    5.1.5 Pupil interview showcards .............................................................................. 104
    5.1.6 Stakeholder interview ...................................................................................... 117
    5.1.7 Parent focus group discussion guide ................................................................. 121
  5.2 Analysis ..................................................................................................................... 125
    5.2.1 Data preparation record .................................................................................... 125
    5.2.2 Interpreting statistics ....................................................................................... 126
  5.3 Features of well-designed trials .............................................................................. 129

This research has been carried out in compliance with the International standard ISO 20252, (the International Standard for Market and Social research), the Market Research Society’s Code of Conduct and UK Data Protection law.
Acknowledgements

The research team are grateful for the help of the following people, who have given us their time and enabled us to carry out this evaluation:

1. The authors of the Kids Empathic Development Scale who have patiently answered our questions and given us permission to use their scale:

2. Key individuals from local authorities who helped us to make contact with the evaluation schools:
   - Glasgow
   - Highland
   - South Lanarkshire
   - West Dunbartonshire
   - West Lothian

3. Head teachers, class teachers, Roots of Empathy instructors and pupils in the evaluation schools

4. Parents who took part in our focus groups, from evaluation schools from West Dunbartonshire and Glasgow

5. Parents who took part in in-depth interviews from schools in Inverclyde and Glasgow

6. A range of stakeholders who took part in in-depth interviews. Roles included:
   - Class teachers
   - Depute heads
   - Development officer
   - Educational psychologists
   - Head teachers
   - Quality improvement officer
   - Roots of Empathy instructors
   - Roots of Empathy instructor mentor
   - Roots of Empathy programme manager
   - Service managers

7. Staff from Action for Children
Executive summary

Introduction and aims

Previous research studies of Roots of Empathy from around the world have reported significant reductions in aggression and increases in prosocial behaviour. Action for Children and partners from the Early Years Taskforce commissioned an external review of these studies by the Social Research Unit at Dartington to help direct the focus of this current study.

Therefore, Action for Children commissioned a research team (Qa Research and University of Glasgow) to carry out research to examine the extent to which the Roots of Empathy intervention in Scotland works to reduce levels of aggression with 5-8 year old children and how changes in empathy mediate this relationship. Roots of Empathy is a classroom-based programme that aims to reduce levels of aggression within schoolchildren, while increasing their social and emotional competence and helping develop their empathy.

The design of the research

The research team set out to design the most rigorous and robust research method possible. In clinical settings, the gold standard is recognised to be a randomised controlled trial (RCT), but this is not normally feasible or ethical within educational settings.

So, a design known as ‘quasi-experimental’ was adopted, as is typically used in social science settings to evaluate the benefits of an intervention. The ‘strongest’ type was chosen, intentionally to show as much evidence of causation between the Roots of Empathy intervention and the measured outcomes, with the following two key features:

- A control group of classrooms was used to act as a point of comparison to the group of Roots of Empathy (intervention) classrooms
- Data was collected about pupils in both groups of classrooms at two time points
  - Early in the school year, before the start of Roots of Empathy = baseline
  - At the end of the school year, after the Roots of Empathy = follow-up

There are two broad types of data collection that can be used in research trials:

- **Quantitative** = gathers quantities of data that can be measured and analysed numerically, for example via structured surveys or online polls
- **Qualitative** = allows deeper exploration of themes as it gathers opinions, perceptions, motivations and barriers, for example via focus groups and in-depth interviews
This study collected two types of quantitative data at both time points.

1. Teachers completed questionnaires including items from standardised scales to report on pupils’ levels of two different types of empathy (affective, cognitive), aggression and prosocial behaviour.

2. Researchers interviewed a smaller sample of the pupils, using a structured picture-based tool called the Kids Empathic Development Scale (KEDS), to measure three aspects of empathy (affective, cognitive and behavioural).

The study also collected complementary qualitative data at the follow-up stage:

1. In-depth interviews with a range of stakeholders (head teachers, local authority staff, Roots of Empathy instructors and others).

2. Focus groups or in-depth interviews with parents.

3. Verbatim comments from teachers in response to open questions within the questionnaire.

**Numbers involved – achieved sample**

For the teacher questionnaire:
- 31 class teachers completed baseline questionnaire
- 29 class teachers completed follow-up questionnaire (94% of those from baseline)
- Teachers came from 17 different schools
- Schools came from 5 local authority areas - **Glasgow, Highland, South Lanarkshire, West Dunbartonshire, West Lothian**
- 695 pupils covered at baseline
- 661 pupils covered at follow-up

For the pupil KEDS interview:
- 38 classes, one intervention and one control from each school visited
- Classes visited were from 19 schools
- 144 pupils were interviewed at baseline
- 112 pupils were interviewed at follow-up

For the qualitative data collected:
- 24 stakeholders were interviewed, from 25 local authorities
- 13 parents attended focus groups or were interviewed
- 31 teachers provided comments at baseline
- 29 teachers provided comments at follow-up
Significant outcomes

This study found significant results showing that the Roots of Empathy intervention ‘worked’. This is a strong finding, given that it can be difficult to achieve such results in social science research when there are so many different factors taking place in pupils' lives that could have impacts on the outcomes.

The study made a new contribution to the body of evidence, in showing that empathy is a direct outcome from the Roots of Empathy intervention; and that increased empathy led to increased prosocial behaviour.

Findings from teacher questionnaire

Changes in pupils across the school year

Affective empathy is when children feel the same feelings as other people, so for example, feeling sad when another child is sad, or being happy for peers when they are happy.

- When looking at the amount of change in affective empathy and analysing for statistical significance, it was found that pupils' affective empathy significantly increased across the school year in the intervention group, while it did not change significantly in the control group.

Cognitive empathy is the extent to which children understand why other people feel the way they feel, for example, understanding the reasons why another child is crying.

- When looking at the amount of change in cognitive empathy and analysing for statistical significance, it was found that there did not appear to be any changes in pupils' cognitive empathy across the school year, nor between the intervention and control groups. This may be because cognitive empathy was already rated quite high for both groups at the beginning of the year (4.0+ on a scale from 1.0 to 5.0): this is known as a ceiling effect.

Aggression is the extent to which children show hostile or even violent behaviour or attitudes toward other children.

- When looking at the amount of change in aggression and analysing for statistical significance, it was found that pupils' aggression decreased across the school year in intervention classes, while it increased across the school year in control classes.

Prosocial behaviour is the extent to which children act positive, friendly, and helpful towards their peers.

- When looking at the amount of change in prosocial behaviour and analysing for statistical significance, it was found that prosocial behaviour increased across the school year in both the intervention and control samples, although it increased more in the intervention group.
Changes in pupils in need across the school year

To find out whether the intervention worked equally well for those pupils who are most in need, further analysis was carried out on scores from those pupils who started out with the lowest 25% empathy scores or with the highest 25% aggression scores.

For those starting out low in affective empathy:

- These pupils had a greater increase in their affective empathy across the school year in the Roots of Empathy intervention group, than in the control group

For those starting out low in cognitive empathy:

- Unlike the overall cohort of pupils who showed no significant change in cognitive empathy, these pupils did tend to increase in cognitive empathy across the year, and with a greater increase for Roots of Empathy pupils, compared to those in the control group

For those who started out with the highest levels of aggression:

- These pupils in the Roots of Empathy group had a decrease in aggression across the school year, while those who were in the control classrooms increased in aggression

For those starting out low in prosocial behaviour:

- These pupils from both groups increased a little across the year, although for Roots of Empathy pupils this increase was larger than the increase for those in the control group

Differences by gender

Across the school year, boys in the intervention group decreased in aggression, increased in prosocial behaviour, and increased in affective empathy.

Girls in the intervention group decreased in aggression across the school year.

The results suggest that the intervention may be more effective for boys than for girls, on average, although the qualitative findings suggest that girls gained softer outcomes including improved resilience or emotional literacy.

Empathy as a factor influencing changes in prosocial behaviour

Mediation analysis showed that the impact of the intervention on prosocial behaviour occurred because of an increase in affective empathy, which then predicted an increase in prosocial behaviour.
Findings from the pupil interview

Changes in pupils across the school year

Interviewing pupils with KEDS tool was time intensive but valuable as it helped to triangulate the data (data sourced via more than one method) and provided an alternative where the pupil could interact with a researcher and speak for him- or herself.

There were two areas where the analysis of the KEDS data provided similar results to the teacher-report:

- The intervention appears to have a slight advantage in terms of increases in affective empathy compared to no change in affective empathy for the control group

- When examining those pupils who start out most in need (bottom 50% for empathy), there is indication that pupils who are lowest in cognitive empathy benefit from the intervention and increase more in cognitive empathy than pupils in the control.
Qualitative findings

Impact of the programme on pupils’ empathy

Teachers noted examples of the positive effect of the programme on pupils’ affective empathy and cognitive empathy:

“Before, I would have had to say, ‘Can you see so and so is upset?’ Now, they can see that for themselves.” (Teacher)

Impact of the programme on pupils’ aggression

Teachers and stakeholders noted positive effects of the programme on some pupils’ levels of aggression. They gave examples of the pupils’ more constructive approaches to solving differences since taking part in the programme:

“…for some of the girls in particular, that could be a wee bit spoilt, selfish and get involved in disputes….it has made a big difference to them. They’ve taken it on board; they’ve taken a closer look at themselves and they’ve actually been behaving a little bit better with others and it’s changed them.” (Teacher)

Impact of the programme on pupils’ prosocial behaviour

Some teachers, stakeholders and parents had noticed that pupils were showing more caring, supportive and sharing behaviour with peers. Parents attributed this to the programme:

“…the importance of kindness and behaviour…they are very, very aware of these things and I think the programme helped 100% to achieve that.” (Parent)
Other impact on pupils

**Enjoyment**
Stakeholders, parents and researchers observed the pupils’ enthusiasm for and enjoyment of the programme:

“My daughter was crying the day it finished…her saying [baby] is not coming in anymore…she really enjoyed it.” (Parent)

**Emotional literacy**
For teachers and stakeholders, their expectations had been met. They were impressed with pupils’ grasp of emotional concepts and increased, correct use of emotional vocabulary:

“They have an understanding of their own learning and development and communicate feelings freely.” (Teacher)

**Resilience**
Stakeholders and parents gave examples of how the programme had helped some pupils to display resilience in the face of difficulties – death of a father, sibling with leukaemia, hearing impairment and difficult family circumstances:

- One parent recounted how their child improved their communication skills, despite a difficult year for the family; and that this resulted in them being able to play and be more involved with friends, with the child now feeling more confident and popular

**A higher priority for emotional wellbeing**
Stakeholders had noticed a shift in pupils’ priorities away from materialism towards emotional wellbeing:

“They talked about [the baby] being included, as opposed to having a bike or lots of money. It was about being loved and having lots of friends.” (Stakeholder)

**Knowledge sharing**
Stakeholders and parents reported that pupils enjoyed sharing their knowledge about babies, health and wellbeing from the programme with their families:

“They were telling me how to do things. My daughter would play with her dolls and show me!” (Parent)
Impact on other people

**On families of pupils experiencing the programme**
Stakeholders noted that the programme had helped improve the home-school relationship. A parent noticed that it was easier to resolve issues of sibling rivalry at home, due to the programme.

**On other pupils in the school**
Stakeholders and teachers observed a ripple effect of the programme, spreading to other pupils in school – transmitted via the programme noticeboard and via contact in the playground.

**On teachers in the school**
Some teachers and head teachers had learnt new ways to deliver the curriculum, through the programme.

**On the Roots of Empathy baby**
Stakeholders and parents noted the boost to development for the baby.

**Process evaluation – what makes Roots of Empathy work**

**Programme format**
For most stakeholders, the structure of the programme worked well, with the baby visit reinforcing the classroom sessions on the same theme.

**Baby visits**
All groups of participants saw the baby visits as the defining and popular feature of the programme.

**Individual sessions**
Stakeholders appreciated sessions on crying, emotions and safety; parents commented positively on sessions about milestones and communicating.
Conclusions

The evaluation findings have a number of implications for practice regarding the Roots of Empathy intervention. These are our conclusions, based on the results of the evaluation:

**Impact on pupils**

- Pupils receiving the Roots of Empathy programme significantly increased their affective empathy and prosocial behaviour across the school year, compared with control pupils, not receiving the programme. Pupils receiving the programme also decreased in aggression, whilst those in control classes increased in aggression across the school year.

- The Roots of Empathy intervention tends to have its greatest impact for children with the lowest empathy and prosocial skills and who are the most aggressive. The qualitative findings suggested that the programme could bestow increased resilience on pupils who were experiencing personal, familial or health-related challenges.

- Accordingly, when administering Roots of Empathy interventions, the study findings support an approach of targeting classes, schools and local authorities where there is evidence of higher aggression, lower empathy and lower prosocial skills among children (such as a large number of detentions, behavioural problems and incidents, and/or records of emotional and behavioural difficulties, among others).

- Boys tend to benefit more than girls, although girls did also display reductions in aggression. The qualitative findings suggest that girls gained other softer outcomes including improved resilience or emotional literacy.

**Sustained impact on pupils**

- It is beyond the scope of this evaluation to demonstrate whether benefits gained from the programme will be retained by pupils or whether the benefits may even increase if trajectories of change (in empathy, aggression and prosocial behaviour) continue throughout pupils’ development. Qualitative findings indicated the belief in many teachers, stakeholders and parents that the programme would have long-lasting benefits for the participating pupils, but only a long-term evaluation, following the sample a year or more into the future, could provide the answers.
1. **Introduction and aims**

Action for Children commissioned a research team comprising Qa Research and University of Glasgow to carry out:

“…research that will examine the extent to which the Roots of Empathy intervention in Scotland works with 5-8 year old children and how changes in empathy mediate this relationship.”

Action for Children’s invitation to tender document described Roots of Empathy as a classroom-based programme that aims to reduce levels of aggression within schoolchildren, while increasing their social and emotional competence and helping develop their empathy.

The programme involves nine family visits to the classroom over the school year, with a baby as the focus. Within this context, the pupils observe the loving parent-child relationship and may grow to understand the baby’s intentions and emotions. A trained Roots of Empathy instructor guides classroom sessions the week before each visit and the week afterwards each visit.

The Roots of Empathy instructor leads activities designed to help pupils identify and understand:

- Their own feelings
- The feelings of others
- What makes them unique
- How to care for the baby safely

Previous research studies of Roots of Empathy from around the world have reported significant reductions in aggression and increases in prosocial behaviour. Action for Children and partners from the Scottish government Early Years Taskforce commissioned an external review of these studies by the Social Research Unit at Dartington to help direct the focus of this current study, particularly on:

- **Younger primary school aged children**, compared to many previous studies that focused on older primary school aged children

- **The mediating role of empathy** on any impacts noted on pupil aggression or prosocial behaviour

---

1 Action for Children, Roots of Empathy Invitation to Tender, July 2014
2. **Methodology**

2.1 **Designing a robust evaluation methodology**

The research team presented their original proposed method for the evaluation of Roots of Empathy in Scotland to Action for Children on 13th August 2014.

After Action for Children commissioned the evaluation, the organisation provided an evidence review by the Social Research Unit at Dartington to the research team for our consideration when refining the evaluation methodology.

Action for Children also referred the research team to consult the Standards of Evidence Developed by the Social Research Unit at Dartington\(^2\).

The research team consulted other documents that provided guidance on robust evaluation methodology, including:

- UK Magenta Book\(^3\)
- Maryland Scientific Methods Scale\(^4\)
- DIY Evaluation Guide\(^5\)
- Investing in Children: an Overview\(^6\)
- What Counts as Good Evidence\(^7\)
- The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics \(^8\)

The research team took into consideration Action for Children’s stated preference for mainly quantitative methods of data collection and an emphasis on impact rather than process evaluation. The scope and design of the research was also influenced by the financial and time constraints available and the fact that schools and classes had already been selected non-randomly to receive the Roots of Empathy intervention.

The research team set out to design the most rigorous and robust research method possible. In clinical settings, the gold standard is recognised to be a randomised controlled trial (RCT), but this is not normally feasible or ethical within educational settings. Quasi-experimental designs are regarded as ‘a second best choice’ to true experiments (Robson, 1993, p.98\(^9\)) and are commonly used in social science settings to evaluate the benefits of an intervention.

The research team adopted a classical quasi-experimental pre/post design, with the emphasis on quantitative data collected by teacher-report questionnaires at the beginning and end of the school year. Its key features were:

---

\(^2\) The Social Research Unit at Dartington, The ‘What Works’ Standards of Evidence, 2013


\(^4\) David P Farrington, Denise C Gottfredson, Lawrence W Sherman, Brandon C. Welsh, Maryland Scientific Methods Scale, Evidence-Based Crime Prevention, pp 13-21, 2002


\(^6\) The Social Research Unit at Dartington, The ‘What Works’ Standards of Evidence, 2013

\(^7\) Research Unit for Research Utilisation, What Counts as Good Evidence? Provocation Paper for the Alliance for Useful Evidence, 2013


• A control group of classrooms used to act as a point of comparison to the group of Roots of Empathy (intervention) classrooms

• Data was collected about pupils in both groups of classrooms at two time points
  o Early in the school year, before the start of Roots of Empathy = baseline
  o At the end of the school year, after the Roots of Empathy = follow-up

This study collected two types of quantitative data at both time points:

1. Teachers completed questionnaires including items from standardised scales to report on pupils’ levels of two different types of empathy (affective, cognitive), aggression and prosocial behaviour

2. Researchers interviewed a smaller sample of the pupils, using a structured picture-based tool called the Kids Empathic Development Scale (KEDS), to measure three aspects of empathy (affective, cognitive and behavioural)

The study also collected complementary qualitative data at the follow-up stage:

1. In-depth interviews with a range of stakeholders (head teachers, local authority staff, Roots of Empathy instructors and others)

2. Focus groups or in-depth interviews with parents

3. Verbatim comments from teachers in response to open questions within the questionnaire

See Appendix 5.3 for a table showing features of well-designed trials; and how this study compares.

The details of the method follow.
2.2 Sampling and selection of research participants

2.2.1 Sampling of local authorities, schools and classes

Across Scotland, 27 local authorities had some schools that were receiving the Roots of Empathy programme during the school year 2014-15. The research team have not been privy to the selection process by which local authorities, schools and whole classes had been selected to receive the Roots of Empathy. However, it appears that none had been randomly sampled.

Action for Children used purposeful strategic sampling in order to obtain a sample of five local authorities, from the 27 local authorities to be invited to take part in the study: Glasgow, Highland, South Lanarkshire, West Dunbartonshire and West Lothian.

Then, from those five authorities, Action for Children and the authorities sampled 20 schools to be invited to the study, include a mix of urban and rural schools, and to include those schools already demonstrating a strong commitment to Roots of Empathy.

Within each school that was invited and agreed to take part, the study involved two classes:

- Intervention class, receiving the Roots of Empathy programme this year
- Control class, not receiving the Roots of Empathy programme this year

The classes ranged between year groups Primary 1 and Primary 5, with some composite classes containing pupils of more than one school year group, for example, Primary 1 and Primary 2.

The research team intended that the control class should ideally be of a different year group (one lower or one higher) than the intervention class to avoid the potential confounding effect of children from the same year group talking to each other in the playground about what they were learning through the Roots of Empathy programme. However, head teachers retained the final say on the selection of the control class and there were some instances where the intervention and control classes in a school were the same year group.

2.2.2 Sampling for the teacher report questionnaire

Nineteen of the 20 schools decided to engage with the research.

In each of the 19 schools, teachers from both the intervention class and the control class – 38 class teachers - were asked at the beginning of the school year to complete the baseline questionnaire including ratings for all of their current class pupils.

At follow-up, at the end of the school year, only those 31 class teachers who had completed the baseline questionnaire for all or most pupils in their class were invited to complete the questionnaire once again.
2.2.3 Sampling for the pupil interview

At baseline, the team visited all 19 schools involved in the evaluation, with a target to interview the following numbers of pupils:

- 5 interviews with pupils from the intervention class
- 2 interviews with pupils from the control class

The purpose of the sampling rationale was to assess students’ empathy, with a focus on those within the intervention classes, since one-on-one interviews are time consuming and costly. However, to facilitate some comparison, a smaller subset of pupils in the control classes was also included. Class teachers selected pupils for interview, with a request from researchers to provide a range (of ages, gender and ability) and to ensure that the pupils chosen were likely to feel comfortable taking part in the research.

At follow-up, researchers visited all 19 schools and aimed to re-interview the same pupils, if available, to allow tracking of individual pupil changes in empathy.

2.2.4 Selection of stakeholders for in-depth interview

A sample frame was prepared for 24 qualitative in-depth interviews to take place with a range of stakeholders who may have relevant views to contribute about the Roots of Empathy programme.

Roles of potential stakeholders included:

- Class teachers
- Depute heads
- Development officer
- Educational psychologists
- Head teachers
- Local authority link people
- Quality improvement officer
- Roots of Empathy instructors
- Roots of Empathy instructor mentor
- Roots of Empathy programme manager / service co-ordinator
- Action for Children managers
Local authorities included:

- Aberdeenshire
- Angus
- Argyll & Bute
- Clackmannanshire
- Dumfries & Galloway
- Dundee
- East Ayrshire
- East Lothian
- Edinburgh
- Fife
- Glasgow
- Highland
- Inverclyde
- Moray
- North Ayrshire
- North Lanarkshire
- Orkney
- Perth & Kinross
- Scottish Borders
- South Ayrshire
- South Lanarkshire
- Stirling
- West Dunbartonshire
- Western Isles
- West Lothian

2.2.5 Selection of parents for focus group or interview

Researchers liaised with staff from schools in each of the five local authorities directly taking part in the evaluation to gauge their interest in hosting a focus group for parents. With the help of staff, researchers recruited and hosted focus groups with parents of pupils in intervention classes from two schools. Researchers interviewed other parents who wanted to contribute via in-depth telephone interviews.
2.3 Research tool design

All the tools are available in appendix 5.1.

2.3.1 Teacher-report questionnaire (quantitative)

The research team carried out a review of available standardised scales that would be credible and robust to measure empathy, aggression and prosocial behaviour in the following groups of pupils:

- At baseline (pre-intervention) with pupils from intervention classes
- At baseline with pupils from control classes
- At follow-up (post-intervention) with pupils from intervention classes
- At follow-up with pupils from control classes

The research team considered the following factors before making recommendations for the selection of measures:

- Teacher-reported standardised scales
- From peer-reviewed journals
- Appropriate for the age group in the current project – Primary 1, 2, 3 and 4 year groups
- As few items as possible to be manageable for teachers to complete
- Ideally validated in the UK

To measure empathy, the research team chose the two most robust items representing affective empathy and the two most robust items representing cognitive empathy from the Griffith Empathy Measure (Dadds et al., 2008)\(^\text{10}\). This is a validated parent-report scale, which was easily adapted for the teacher-report measure. The scale was previously validated with pupil samples and is appropriate for ages 4-16.

These items represent affective empathy:

- This child gets upset when another person is acting upset
- This child cries or gets upset when seeing another child cry

These items represent cognitive empathy and were reverse coded:

- This child can’t understand why other people get upset
- This child rarely understands why other people cry

To measure *aggression*, the research team chose the four items from the overt aggression variable in Crick’s (1996) Pupil’s Social Behaviour Scale-Teacher:\(^{11}\):

- This child hits, shoves, or pushes peers
- This child initiates or gets into physical fights with peers
- This child threatens to hit or to beat up other children
- This child tries to dominate or bully peers

To measure *prosocial behaviour*, the research team chose the four items from the prosocial subscale in Crick (1996) Pupil’s Social Behaviour Scale-Teacher:

- This child says supportive things to peers
- This child tries to cheer up peers when they are sad or upset about something
- This child is helpful to peers
- This child is kind to peers

Items were presented in a random order, to ensure that there were not order effects (a form of bias that can be due to practice effects and/or fatigue effects) when completing the questionnaire.

Teachers were asked to rate each pupil for each item, on a scale as follows:

- From 1 = "never true of this pupil"
- To 5 = "always true of this pupil"

The research team used the above measures to form the basis of a teacher questionnaire that they set up on a Computer Aided Web Interviewing (CAWI) system.

The questionnaire also asked teachers to supply details about their school / class / pupils and to answer a free text questions about their expectations for the Roots of Empathy programme.

Example screens are shown below:

The follow-up questionnaire was quicker and easier for teachers to complete as it imported (in random order) pupil names, rather than requiring them to be typed. A few free text questions were included about impact of the programme, how the programme met expectations and any suggested improvements to programme delivery.

See appendices 5.1.1 and 5.1.2 for baseline and follow-up versions of the teacher questionnaire.

2.3.2 Pupil interview (quantitative)

Action for Children had indicated they preferred a primarily quantitative approach to the research, so the research team obtained permission from the developers to use the Kids Empathic Development Scale (KEDS) tool\(^\text{12}\), developed by Reid et al. (2013), for face-to-face interviews with pupils to measure three types of empathy:

- Affective empathy
- Cognitive empathy
- Behavioural empathy

The measure had been developed for young children who cannot yet appropriately report their own levels of empathy.

This tool provided more objectivity to the evaluation by allowing external researchers to conduct part of the assessment rather than rely solely on teacher-report measures, which may have some level of bias. KEDS data and teacher-report data were included together to provide a complementary and comprehensive assessment of empathy.

Using a very structured interview process, the tool assesses pupil responses to picture scenarios depicting a range of individual and inter-personal situations, differing in social complexity. Pupils came out of class one by one to take part in the KEDS interview in a quiet place away from classmates.

First, a screener exercise – the Emotional Perspectives Card - is administered, to ensure that pupils who proceed with the interview can identify at least the emotions sad, happy and angry from a set of cartoon faces.

![Emotional Perspectives Card](image)

Twelve picture scenarios are then shown in turn and a standardised questioning format is applied for each.

To assess **affective empathy**, the following question is asked:

- How do you think this [girl / boy / man] feels? Pick one of the faces on the card that best matches how she feels

To assess **cognitive empathy**, the following question is asked, with the follow-up prompt:

- Can you tell me why this [girl / boy / man] feels ……?
- Can you tell me more about what is happening in this picture?

---

To assess **behavioural empathy**, the following question is asked:

- What would you do if you were that [girl / boy / man]?

These responses are then scored, according to a systematic scoring protocol, for **affective**, **cognitive** and **behavioural empathy** components.

See appendices 5.1.3 and 5.1.4 for baseline and follow-up versions of the pupil interview script. See appendix 5.1.5 for pupil interview showcards.

### 2.3.3 Stakeholder interview (qualitative)

A semi-structured interview guide was developed, to be used with a range of stakeholders and covering the following themes:

- Background
- Delivery
- Expectations
- Impact and outcomes
- Parental engagement

See appendix 5.1.6 for stakeholder interview guide.

### 2.3.4 Parent focus group or interview (qualitative)

A discussion guide was developed which included a participatory appraisal techniques, so that all parents could feel comfortable to take part. Themes covered included:

- Awareness of the Roots of Empathy programme
- Perceptions of the programme, gained from their child or from school
- Perceptions of positive or negative impact of the programme
- Suggestions for developing the programme, including involving parents

See appendix 5.1.7 for parent interview guide / parent focus group discussion guide.
2.4 Data collection

2.4.1 Baseline teacher questionnaire

Action for Children liaised with key contacts from each local authority to obtain school contact details, before passing them to the research team, who invited teachers to engage with the questionnaire in the autumn term of 2014.

The research team used the school contact details provided to send an email to schools containing:

- A full explanation of the evaluation and details of how schools were invited to engage
- A parent letter template that schools could tailor and send to parents so that parents were aware of the research and could withdraw consent if wished
- A web link to the online questionnaire for the teachers to complete

Teachers were asked to answer the baseline questionnaire before any Roots of Empathy activity had taken place in the school.

The research team provided regular updates to Action for Children and the local authority contacts in order that they could help prompt or support teachers to complete the questionnaire. The team also re-contacted schools via email or telephone to try to achieve the maximum number of completions.

Teachers completed questionnaires over a two-month period, between 25th September and 21st November 2014.

2.4.2 Follow-up teacher questionnaire

In May 2015, the research team emailed schools with the web link for the online questionnaire and a copy of the parent letter. The team also informed schools that Action for Children had decided to provide a gift voucher direct to teachers after their completion of the follow-up questionnaire.

The team provided regular updates to Action for Children and the local authority contacts in order that they could help prompt or support teachers to complete the questionnaire for a couple of weeks into June, if needed. The team also re-contacted schools via email or telephone to try to achieve the maximum number of completed questionnaires.

Teachers completed questionnaires over roughly a one-month period at the end of the school year, between 19th May and 26th June 2015. Action for Children issued a £15 gift voucher to each teacher completing a questionnaire.
2.4.3 Baseline pupil interview

The timing of the baseline data collection via pupil interview was also considered, so that pupils would ideally be interviewed before they could have received any potential benefits from the Roots of Empathy programme.

The first theme of the programme, Theme One, consists in successive weeks of three sessions:

1. Pre-family visit to the class by the Roots of Empathy instructor
2. Family visit to the class by parent and baby
3. Post-family visit to the class by the Roots of Empathy instructor

Action for Children reported that it is only at the Theme One post-family visit that the pupils’ learning about feelings and perspective-taking may start to be integrated. Therefore, it was decided that pupil interview data would ideally be collected before the Theme One post-family visit (session three).

Interviews took place over a one-month period, from 1st to 30th October 2014.

The research team drafted a briefing document for the administration of the KEDS tool, to ensure that interviews were carried out consistently and according to the directions of the author of KEDS. Interviewers worked with schools to cause the minimal amount of disruption to the school day, whilst still obtaining their target number of interviews, where possible.

2.4.4 Follow-up pupil interview

The dates for final Roots of Empathy sessions within schools was not known as many schools were continuing to re-schedule these even at a late stage in the summer term. Therefore, the research team aimed to complete follow-up pupil interviews as close as possible to the end of the summer term 2015. They took place over a one-month period at the end of the school year, from 28th May to 28th June 2015.

The lead researcher re-briefed interviewers before the follow-up phase. Where possible, the same interviewers were used.

The research team worked with schools to cause the minimal amount of disruption to the school day, whilst still obtaining their target number of interviews, where possible.

2.4.5 Follow-up stakeholder interview

An email was sent to potential stakeholders in May, explaining the research and inviting them to take part in an in-depth interview.

24 interviews were booked in and carried out by telephone over a six-week period from 1st June to 15th July, at a time of day to suit the participants.
2.4.6 Follow-up parent focus group and in-depth interview

The research team made contact with schools to ask if they would be interested in hosting a focus group with parents from their intervention class, either at lunchtime or after school.

Interested schools issued a letter to parents to explain the research and invite them to attend a focus group, with the promise of a £20 gift voucher for those who attended.

Researchers moderated two focus groups (with six and four participants respectively) during lunch times, alongside the school visits to complete pupil interviews, both on 15th June 2015.

Additional parental views were captured during a 12-day period, from 19th June to 1st July, via five in-depth telephone interviews with parents from evaluation schools and from those in different local authorities, outside the evaluation. Interviewees received a £10 gift voucher.
3. Key findings

Key findings follow from each wave of data collection:

- Quantitative findings from teacher questionnaire
- Quantitative findings from administering KEDS protocol with pupils
- Qualitative findings from class teachers, stakeholders and parents

The findings are then organised under topic headings, with summary findings boxed under each heading; and detailed findings will follow.

Summary findings will be highlighted like this.

3.1 Quantitative findings from teacher questionnaire

This section is examining the impact of the Roots of Empathy programme by analysing changes in the four key variables from pre- (beginning of the school year) to post- (end of school year), primarily analysing differences between intervention (Roots of Empathy) classes and control classes and including analysing differences for sub-groups (gender, age, bottom quartile, classroom).

Please note that graphs in this section do not always show the full range of values for each variable; for example, the values on the x-axis do not always start at zero. This allows graphs to highlight and draw attention to important differences in the data. Therefore, when interpreting the graphs, special attention should be paid to the values on the x-axis of each graph.

For details of the data preparation record and explanations of some of the statistical tests used on the quantitative date, please see appendix 5.2.
3.1.1 Achieved sample

In total, 31 class teachers participated in the baseline questionnaire and 29 class teachers took part in the follow-up questionnaire (94% of those from baseline).

Respondents came from 17 different schools across the five different local authority regions.

Data was provided by class teachers for 695 pupils at the pre-test and 661 pupils at the post-test.

Out of the 20 schools originally expected to take part in the evaluation, one chose not to engage at all.

Therefore, the research team invited 19 schools (with two class teachers in each) to complete the questionnaire (38 teachers).

In total, 31 class teachers participated in the baseline questionnaire and 29 class teachers took part in the follow-up questionnaire (94% of those from baseline).

Respondents came from 17 different schools across the five different local authority regions.

Data was provided by class teachers for 695 pupils at the pre-test and 661 pupils at the post-test.

Therefore, there were 34 pupils for whom data was provided only at pre-test. Two class teachers from control classes did not participate in the questionnaire at post-test. Therefore, teacher-report data was provided for 661 pupils for both pre- and post-test. This corresponds to a 4.9% attrition rate across the study.

Of these 661 pupils, there were 309 pupils in the control group and 352 pupils in the intervention group.

3.1.2 Reliability of scores

The reliability of the scale and a measure of its internal consistency was assessed, using Cronbach's alpha.

The Cronbach’s alpha scores for each scale were good (0.9 > α ≥ 0.8) or excellent (α ≥ 0.9), indicating that the measures were reliable:

- For affective empathy, α = .84 and .89 at pre and post
- For cognitive empathy (items reverse coded), α = .80 and .81 at pre and post
- For aggression, α = .92 and .91 at pre and post
- For prosocial behaviour, α = .91 and .92 at pre and post

Given the reliability of these adapted measures, combined with randomised implementation, the tool’s usability for future studies is worth considering.
3.1.3 Changes in the four variables from pre- to post-, by intervention and control groups

The four main variables (affective empathy, cognitive empathy, aggression, and prosocial behaviour), at both pre- and post-test, were calculated by taking the average (mean) of the items within each scale, together with standard deviation.

The purpose of multiple-item scales is to reduce measurement error and increase reliability of the variable. Possible scores ranged from 1-5, where 1 means low levels of this trait (never true) and 5 means high levels of this trait (always true).

Here are the overall pre- and post- mean scores and standard deviation for the four variables:

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>SD</th>
<th>Post</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTERVENTION (n = 352)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affective Empathy</td>
<td>2.19</td>
<td>.93</td>
<td>2.43</td>
<td>1.12</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
<td>4.17</td>
<td>.89</td>
<td>4.26</td>
<td>.90</td>
</tr>
<tr>
<td>Aggression</td>
<td>1.63</td>
<td>.96</td>
<td>1.53</td>
<td>.87</td>
</tr>
<tr>
<td>Prosocial Behaviour</td>
<td>3.82</td>
<td>.80</td>
<td>4.11</td>
<td>.78</td>
</tr>
<tr>
<td><strong>CONTROL (n = 309)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affective Empathy</td>
<td>1.85</td>
<td>.79</td>
<td>1.84</td>
<td>.85</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
<td>4.17</td>
<td>.81</td>
<td>4.24</td>
<td>.83</td>
</tr>
<tr>
<td>Aggression</td>
<td>1.50</td>
<td>.79</td>
<td>1.66</td>
<td>.93</td>
</tr>
<tr>
<td>Prosocial Behaviour</td>
<td>3.67</td>
<td>.89</td>
<td>3.81</td>
<td>.99</td>
</tr>
<tr>
<td><strong>TOTAL (n = 661)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affective Empathy</td>
<td>2.03</td>
<td>.88</td>
<td>2.15</td>
<td>1.04</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
<td>4.17</td>
<td>.85</td>
<td>4.25</td>
<td>.87</td>
</tr>
<tr>
<td>Aggression</td>
<td>1.57</td>
<td>.89</td>
<td>1.59</td>
<td>.90</td>
</tr>
<tr>
<td>Prosocial Behaviour</td>
<td>3.75</td>
<td>.85</td>
<td>3.97</td>
<td>.90</td>
</tr>
</tbody>
</table>

Note. M = Mean (average); SD = Standard Deviation; n = sample size
This set of results examines whether changes in empathy (affective and cognitive), aggression and prosocial behaviour occurred across the school year and whether there were differences in these changes depending upon whether the pupil was in the Roots of Empathy intervention classroom compared with a control classroom. In order to analyse this data, pupils were split into control and intervention groups and dependent samples t-tests were used to examine within-pupil changes from pre- to post-test.

Next, researchers used general linear models with the whole sample with a within-subject factor (pre to post), a between-subject factor (intervention vs. control), and the interaction between the two, in order to examine whether pupils showed different trajectories of change in their empathy, prosocial behaviour, and aggression depending upon whether they received the intervention or not.

Analysis of baseline data showed that there were significant differences between intervention and control group classrooms at the pre-test in terms of aggression (t(693) = 2.03, p < .05), prosocial behaviour (t(693) = 2.34, p < .05), and affective empathy (t(693) = 5.18, p < .001). There was no significant difference in cognitive empathy at pre-test (t(693) = -.07, ns). Therefore, in general, the mean score differences at post-test between intervention and control group pupils are irrelevant because of these pre-test differences.

In order to account for these differences between the intervention and control groups at baseline, it is important to examine within-pupil changes rather than mean level differences. Therefore researchers did not consider the level the pupil started the year in terms of the teacher’s rating; rather, they analysed whether the teacher rated the pupil as increasing or decreasing across the year. It is also important to consider the interaction term within the analysis model that examines whether the within-group changes from pre- to post-test are different for intervention vs. control groups. This interaction is the key to determining whether the intervention “worked” or not.

Empathy is disaggregated into affective empathy and cognitive empathy, so they are presented separately. This is then followed by results about aggression and prosocial behaviour.
**Affective Empathy**

When looking at the amount of change in affective empathy and analysing for statistical significance, it was found that pupils’ affective empathy significantly increased across the school year in the intervention group, while it did not change significantly in the control group.

Affective empathy is the extent to which children feel the same feelings as others, so for example, getting sad when another child is sad, or being happy for peers when they are happy.

When looking at the amount of change in affective empathy and analysing for statistical significance, it was found that pupils’ affective empathy significantly increased across the school year in the intervention group, while it did not change significantly in the control group. Pupils in the Roots of Empathy intervention group significantly increased in affective empathy from pre- to post-test across the school year, $t(351) = -3.89$, $p < .001$.

Pupils in the control group did not have a change in affective empathy across the school year although the trend suggests a very slight decrease, $t(308) = .54$, ns.

The linear model determined that there was a significant interaction ($F = 11.24$, $p < .001$, $\eta^2 = .02$), meaning that the average change in affective empathy across the school year was significantly different for the intervention group vs. the control group. In other words, the intervention group was increasing while the control group was remaining constant.

The trends are shown below.
**Cognitive Empathy**

When looking at the amount of change in cognitive empathy and analysing for statistical significance, it was found that there did not appear to be any changes in pupils’ cognitive empathy across the school year, nor between the intervention and control groups. This may be because cognitive empathy was already rated quite high for both groups at the beginning of the year (4.0+ on a scale from 1.0 to 5.0): this is known as a ceiling effect in statistics.

Cognitive empathy is the extent to which children understand why other people feel the way they feel, for example, understanding why another child is crying.

When looking at the amount of change in cognitive empathy and analysing for statistical significance, it was found that pupils in the intervention group did not significantly change in their level of cognitive empathy from pre- to post-test across the school year although the trend suggested a slight increase, $t(351) = -1.05$, ns.

Pupils in the control group had a similar pattern, where the level of cognitive empathy did not significantly change from pre to post, $t(308) = -1.48$, ns.

The linear model determined that there was not a significant interaction ($F = .00$, ns), meaning that the average change in cognitive empathy across the school year was the same for both the intervention and control groups. This may be because cognitive empathy was already rated quite high for both groups at the beginning of the year (> 4.0 on a scale from 1.0 to 5.0). This is known as a ceiling effect in statistics, whereby it appears that the independent variable has no effect on the dependent variable because scores were bunched at the upper level of the survey instrument.

The trends are shown below.
**Aggression**

When looking at the amount of change in aggression and analysing for statistical significance, it was found that pupils’ aggression **decreased** across the school year in intervention classes, while it **increased** across the school year in control classes.

Aggression is the extent to which children show hostile or even violent behaviour or attitudes toward other children.

When looking at the amount of change in aggression and analysing for statistical significance, it was found that pupils in the Roots of Empathy intervention group significantly decreased in aggression from pre- to post-test, \( t(351) = 2.33, p < .05 \).

Pupils in the control group significantly increased in aggression across the school year from pre- to post-test, \( t(308) = -4.88, p < .001 \).

The linear model determined that there was a significant interaction \( (F = 25.24, p < .001, \eta^2 = .04) \), meaning that the average change in aggression across the school year was different for the intervention group vs. the control group.

The trends are shown below.

Analysis discovered that pupils’ aggression does not follow a ‘normal’ distribution. Aggression was positively skewed, with most students rated as having low aggression. While the statistical tests are somewhat robust against slight deviations to normality, the deviations were extreme in this case.

Given that the aggression variable was not normally distributed, researchers recommend running nonparametric tests instead of linear tests. The non-parametric test (Wilcoxon Signed Rank) found that the intervention group significantly decreased in aggression \( (p < .05) \) and the control group significantly increased in aggression \( (p < .001) \).
Furthermore, the non-parametric Mann-Whitney U Test was conducted to examine whether the change in aggression from pre- to post-test was significantly different between intervention and control groups, and the test was highly significant at $p < .001$. This confirmed the results above that the two groups differed in their changes in aggression across the school year.
**Prosocial Behaviour**

When looking at the amount of change in prosocial behaviour and analysing for statistical significance, it was found that prosocial behaviour increased across the school year in both the intervention and control samples, although it increased more in the intervention group.

Prosocial behaviour is the extent to which children act positive, friendly, and helpful towards their peers.

When looking at the amount of change in prosocial behaviour and analysing for statistical significance, it was found that pupils in the intervention group significantly increased in prosocial behaviour from pre- to post-test, $t(351) = -6.23, p < .001$.

Pupils in the control group also significantly increased in prosocial behaviour from pre- to post-test, $t(308) = -2.82, p < .01$.

The linear model determined that there was a significant interaction ($F = 5.74, p < .05, \eta^2 = .01$), meaning that the average change in prosocial behaviour was different for the intervention group vs. the control group. Looking at the change scores indicated that pupils in the intervention increased significantly more in their prosocial behaviour than pupils in the control group.

The trends are shown below.

![Change in Prosocial Behaviour from Pre to Post](image)
Summary – changes in the four variables from pre- to post-

The results in this set suggest that Roots of Empathy intervention was effective, as reported by teacher ratings of pupils' empathy, aggression, and prosocial behaviour.

Taking the teachers’ perspective, positive increases in children’s affective empathy and prosocial behaviour as well as a decrease in aggression were observed with the intervention group.

Differences between intervention and control groups were supported statistically, except with children’s cognitive empathy, where no difference was observed.

- Pupils’ affective empathy significantly increased across the school year in the intervention group, while it did not change in the control group.
- Cognitive empathy seemed to change equally, increasing slightly, for both intervention and control pupils.
- Pupils' aggression increased across the school year in the control classes and decreased across the school year in intervention classes.
- Prosocial behaviour increased across the school year in both the intervention and control samples, although it increased more in the intervention group.

Changes in the average scores from pre- to post-test appear small, but these are statistically significant (see Interpreting Statistics section 5.2.2).

The partial eta squared values suggest, on a statistical scale:

- Small effect size for affective empathy
- Small effect size for prosocial behaviour
- Small / medium effect size for aggression

On a more practical scale, they are also meaningful if you consider that there was likely variation between classrooms in the quality or extent to which the Roots of Empathy intervention was implemented as well as variation within classrooms in terms of pupils’ differing development across the year.

The significant interactions from the analysis demonstrate that on average the entire group of intervention pupils significantly improved compared to the entire group of control pupils. If you imagine that the Roots of Empathy intervention were to be implemented perfectly, and if these trajectories were compounded year after year, results would likely be even more substantial.
3.1.4 Examining pupils in need: the bottom 25%

Overall, results suggest that the group of pupils who tend to benefit the most from Roots of Empathy interventions are those who need it the most: those who are low in empathy, low in prosocial behaviour or high in aggression.

The intervention seemed to have no impact on those who were already high in empathy, high in prosocial behaviour, and low in aggression. One way of illustrating this finding is to use the “patient-cure” analogy, whereby medicine given to patients has its maximum impact on those who require it.

A question that naturally arises when evaluating interventions is whether the intervention works equally well for those pupils who are most in need, for example, those pupils who start out with the lowest empathy or the highest levels of aggression.

In order to assess this, researchers compared whether the change from pre- to post-test was greater for those who started out the school year in the bottom quartile compared to those in the middle 50% and in the top quartile.

Quartiles were calculated for each variable based on the initial baseline data for the full sample in order to determine cut-off scores for 25th and 75th percentiles. Using these cut-off scores, pupils were assigned to a group (low, medium, high).

For example, for affective empathy, the 25th percentile fell at a score of 1 (on a 5-point scale) and the 75th percentile fell at a score of 2.5 (on a 5-point scale). Therefore, for affective empathy, pupils with scores of 1 or lower were placed into the “low” group, pupils with scores from 1.1 to 2.4 were placed in the “medium” group, and pupils with scores of 2.5 or higher were placed in the “high” group.

Due to the small range of scores, the cutoff criteria resulted in final group sizes that are not perfectly equal in size.

A series of 3 x 2 ANOVAs was conducted to determine whether there were significant differences in change in affective empathy, cognitive empathy, aggression and prosocial behaviour, based on initial affective empathy scores (low, medium, high) and intervention group (intervention vs. control).

The number of pupils who fell into each quartile group and the mean scores for each group at the baseline data collection are shown in the table below, with sample sizes and mean scores.

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th></th>
<th>Medium</th>
<th></th>
<th>High</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>M</td>
<td>n</td>
<td>M</td>
<td>n</td>
<td>M</td>
</tr>
<tr>
<td>Affective Empathy</td>
<td>186</td>
<td>1</td>
<td>271</td>
<td>1.83</td>
<td>238</td>
<td>3.05</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
<td>179</td>
<td>2.98</td>
<td>260</td>
<td>4.17</td>
<td>256</td>
<td>5</td>
</tr>
<tr>
<td>Aggression</td>
<td>363</td>
<td>1</td>
<td>128</td>
<td>1.35</td>
<td>204</td>
<td>2.73</td>
</tr>
<tr>
<td>Prosocial</td>
<td>188</td>
<td>2.7</td>
<td>331</td>
<td>3.76</td>
<td>176</td>
<td>4.86</td>
</tr>
</tbody>
</table>

Note. n = sample size, M = Mean (average)
Affective empathy

Focusing on those starting out low in affective empathy, pupils had a greater increase in affective empathy across the school year in the intervention group, than in the control group.

Examining significant differences in change in affective empathy, there was a significant difference between intervention and control, with the intervention group having a greater increase in affective empathy than the control group ($F = 26.50, p < .001$).

There was also a significant difference between the low, medium, and high groups ($F = 51.51, p < .001$) and the Tukey post hoc suggested the change in affective empathy from pre to post-test was significantly different for all three groups, with the low group increasing more from pre- to post-test than the medium and high groups.

There was not a significant interaction between the two ($F = .63, ns$), suggesting that effects of being in intervention or control do not differ based on whether the pupil started out in the low, medium, or high groups.

Therefore, the researchers can conclude that the intervention is equally effective in improving affective empathy for all pupils, regardless of pupils’ initial level of affective empathy. For those who start out high in affective empathy, the intervention leads to less of a decrease over the year than for those who are in control, while for those who start out low in affective empathy, the intervention leads to more of an increase over the year than for those in the control.
Pupils who were in the low group (based on bottom quartile cutoff) in terms of their initial affective empathy increased by .88 if they received the intervention while those in the control only increased by .39, a significant difference ($t = 4.31, p < .001$).

Therefore, for pupils who start out with the lowest levels of affective empathy, they all seem to increase a little across the year, although being in the intervention led to an increase that was 2.3 times greater than the increase for those in the control group.

**Cognitive empathy**

Unlike the overall cohort of pupils who showed no significant change in cognitive empathy, these pupils who started out low in cognitive empathy tended to increase in cognitive empathy across the year, with a greater increase for Roots of Empathy pupils, compared to those in the control group.

When examining changes in cognitive empathy, there was not a significant difference between intervention and control ($F = 1.41, p = .24$), which suggests that pupils in the intervention and in the control classrooms increased equally in cognitive empathy across the school year.

There was a significant difference between the low, medium, and high groups ($F = 109.99, p < .001$) and the Tukey post hoc shows that the overall change in cognitive empathy from pre- to post-test was significantly different for all three groups.

Pupils who were initially lowest in cognitive empathy increased the most, the medium group increased slightly, and the high group decreased in cognitive empathy.

Interestingly, there was a significant interaction between the two factors ($F = 12.86, p < .001$), suggesting that effects of being in intervention or control classes differ based on whether the pupil started out in with low, medium, or high levels of cognitive empathy. In other words, the intervention does not impact all pupils equally depending upon their initial levels of cognitive empathy.

As shown in the next figure, the beneficial trend reverses for pupils who started out high, and in fact, being in the intervention leads to a greater decrease than being in the control for these pupils. However, this is likely due to the top quartile of pupils already being rated so high to begin with ($M = 5.0$ on a 5.0 scale), called a ceiling effect. In colloquial terms, these pupils have nowhere to go except for down.
Pupils who were in the low group (based on bottom quartile cutoff scores) in terms of their initial cognitive empathy increased by .94 if they received the intervention while those in the control only increased by .53, a significant difference ($t = 2.58, p < .05$).

Therefore, all pupils who start out low in cognitive empathy tended to increase in cognitive empathy across the year, although being in the intervention group led to an increase that was 1.8 times greater than the increase for those in the control group.

**Aggression**

For pupils who started out with the most problematic levels of aggression, their receipt of the intervention led to a decrease in aggression while those who were in the control classrooms increased in aggression.

When examining changes in aggression scores, there was a significant difference between intervention and control ($F = 33.24, p < .001$), and the mean scores demonstrate that pupils in the control increased in aggression while those in the intervention decreased in aggression.

There was also a significant difference between the low, medium, and high groups in their overall change across the year ($F = 26.14, p < .001$). The Tukey post hoc suggested the change in aggression from pre- to post-test was significantly different for pupils who started out with initially high levels of aggression compared to the medium and low levels, although the medium and low groups did not differ.
Overall, pupils in the low and medium groups increased in aggression, while pupils in the high group decreased in aggression. There was also a significant interaction between the two factors ($F = 13.22, p < .001$), suggesting that effects of being in intervention or control differ based on whether the pupil started out low, medium, or high.

As shown in the figure below, those pupils who started out with high levels of aggression saw the most extreme benefits from the intervention and in fact reversed the direction of their aggression. Those in the medium group still increased in aggression but to less of an extent than those in the control group, while those in the low group had very minimal benefits from being in the intervention over the control.

Here, the meaningful subgroup is the high group (based on top quartile cutoff scores), because the highest levels of aggression are the most problematic.

Pupils who were in the high group in terms of their initial aggression decreased by .50 if they received the intervention and increased by .08 if they were in the control, a significant difference ($t = -4.64, p < .001$).

Therefore, for pupils who start out with the most problematic levels of aggression, being in the intervention led to a decrease in aggression while those who were in the control classrooms increased in aggression.
Prosocial behavior

All pupils who started out with the lowest levels of prosocial behaviour increased a little across the year, although being in the intervention led to an increase that was larger than the increase for those in the control group.

When examining changes in prosocial behaviour scores, there was a significant difference between intervention and control ($F = 14.88, p < .001$), and the mean scores demonstrate that pupils in the intervention increased in prosocial behaviour more than pupils in the control group.

There was also a significant difference between the low, medium, and high groups in their overall change across the year ($F = 47.71, p < .001$) and the Tukey post hoc suggested that the change in prosocial behaviour from pre- to post-test was significantly different between all three groups.

The lowest group increased a lot, the medium group increased less so, and the group that started out high decreased.

There was also a significant interaction between the two factors ($F = 11.78, p < .001$), suggesting that effects of being in intervention or control differ based on whether the pupil started out low, medium, or high.

As shown in the next figure, the results are somewhat similar to those for aggression. Those in the low group for initial prosocial behaviour had a much greater increase in prosocial behaviour when they were in intervention classes compared to those who were in the control classes. Pupils who were already medium or high in their prosocial behaviour did not have much difference between intervention and control. In other words, for pupils who need the intervention the most, the intervention was more effective.
Pupils who were low (based on bottom quartile scores) in terms of their initial prosocial behaviour increased by .83 if they received the intervention while those in the control only increased by .21, a significant difference ($t = 5.20, p < .001$).

Therefore, for pupils who start out with the lowest levels of prosocial behaviour, they all seem to increase a little across the year, although being in the intervention led to an increase that was four times greater than the increase for those in the control group.

**Summary – examining pupils in need**

The overall findings for examining pupils in need (the bottom 25%) suggest that the groups of pupils who tend to benefit the most from Roots of Empathy interventions are those who need it the most. The researchers can speculate that these may include pupils who have disciplinary issues and therefore, often get into trouble with their teachers and other pupils, and/or perhaps also those who might have emotional and behavioural difficulties. With limited resources, it may be advised that these are the pupils who should be given the intervention.

The findings for cognitive empathy, for aggression, and for prosocial behaviour strongly suggest that the groups of pupils who tend to benefit the most are those who are low in empathy, low in prosocial behaviour or high in aggression.

The intervention seemed to have no impact on those who were already high in empathy, high in prosocial behaviour, and low in aggression. For these pupils, there was little room left for them to improve – known as the ceiling effect. Another way of illustrating this finding is to use the “patient-cure” analogy, whereby medicine given to patients has its maximum impact on those who require it, whilst those in good health may not show any health gains after taking the medicine.13

---

13 Please see 3.3 ‘Qualitative Findings’ for other softer outcomes gained by ROE pupils
### Differences by gender

Across the school year, boys in the intervention group decreased in aggression, increased in prosocial behaviour, and increased in affective empathy.

*Girls* in the intervention group decreased in aggression across the school year, with no other significant changes.

The results suggest that the intervention may be more effective for boys than for girls, on average.

One area of particular interest is whether the intervention was more effective for boys or girls, or if the intervention was equally effective regardless of gender.

First, the chart below illustrates the mean scores for girls and boys at pre-test.
In order to determine the effectiveness of the intervention by gender, an independent samples t-test was run separately for boys and girls, with the independent variable of intervention vs. control and the dependent variables of change in aggression, change in prosocial behaviour, change in cognitive empathy, and change in affective empathy.

Here are the pre- and post- mean scores by gender. According to independent samples t-tests, the differences in ratings between girls and boys were significant for all of the variables at both pre- and post-.

<table>
<thead>
<tr>
<th></th>
<th>GIRLS</th>
<th></th>
<th>BOYS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 164</td>
<td>n = 160</td>
<td>n = 203</td>
<td>n = 192</td>
</tr>
<tr>
<td><strong>INTERVENTION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affective Empathy</td>
<td>2.25</td>
<td>.92</td>
<td>2.42</td>
<td>.14</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
<td>4.43</td>
<td>77</td>
<td>4.45</td>
<td>.85</td>
</tr>
<tr>
<td>Aggression</td>
<td>1.38</td>
<td>.66</td>
<td>1.32</td>
<td>.62</td>
</tr>
<tr>
<td>Prosocial Behaviour</td>
<td>4.02</td>
<td>.80</td>
<td>4.30</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CONTROL</strong></td>
<td>n = 153</td>
<td>n = 144</td>
<td>n = 175</td>
<td>n = 165</td>
</tr>
<tr>
<td>Affective Empathy</td>
<td>2.00</td>
<td>.87</td>
<td>2.06</td>
<td>.94</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
<td>4.27</td>
<td>.76</td>
<td>4.27</td>
<td>.76</td>
</tr>
<tr>
<td>Aggression</td>
<td>1.28</td>
<td>.55</td>
<td>1.45</td>
<td>.74</td>
</tr>
<tr>
<td>Prosocial Behaviour</td>
<td>3.88</td>
<td>.85</td>
<td>4.05</td>
<td>.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>n = 317</td>
<td>n = 304</td>
<td>n = 378</td>
<td>n = 357</td>
</tr>
<tr>
<td>Affective Empathy</td>
<td>2.13</td>
<td>.90</td>
<td>2.25</td>
<td>1.07</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
<td>4.35</td>
<td>.77</td>
<td>4.37</td>
<td>.82</td>
</tr>
<tr>
<td>Aggression</td>
<td>1.33</td>
<td>.61</td>
<td>1.38</td>
<td>.68</td>
</tr>
<tr>
<td>Prosocial Behaviour</td>
<td>3.95</td>
<td>.82</td>
<td>4.18</td>
<td>.83</td>
</tr>
</tbody>
</table>

Note: M = Mean (average); SD = Standard Deviation; n = sample size

Independent samples t-test showed that for boys, there was a significant difference between intervention and control on change in aggression, \( t = -3.61 \), \( p < .001 \), a marginally significant difference between intervention and control on their change in prosocial behaviour, \( t = 1.94 \), \( p = .05 \), and a significant difference between intervention and control on their change in affective empathy, \( t = 3.74 \), \( p < .001 \).

The mean scores indicate that boys in the intervention decreased in aggression, increased in prosocial behaviour, and increased in affective empathy.

For girls, there was only one significant difference between intervention and control, and that was on their change in aggression, \( t = -3.66 \), \( p < .001 \). The mean scores indicate that girls in the intervention decreased in aggression.

While the mean scores suggest a general trend that girls also increased in prosocial behaviours and in affective empathy more in the intervention than they did in the control, the differences were not significant within this group.
The gender differences in changes by control and intervention groups are displayed visually below.

**Change in Affective Empathy from Pre to Post**

- Boys: Control > Intervention
- Girls: Control < Intervention

**Change in Cognitive Empathy from Pre to Post**

- Boys: Control > Intervention
- Girls: Control < Intervention
Differences between boys and girls were evident.

There was a significant change for boys who took part in the intervention in terms of aggression, prosocial behaviour, and affective empathy when compared to their control counterparts.

As for girls, improvement only took place in aggression component.

The results suggest that the intervention may be more effective for boys than for girls, on average, since boys have more extreme differences between control and intervention. This makes sense, given that girls start out significantly higher than boys in empathy, prosocial behaviour, and aggression. An alternative explanation could be that boys and girls do not display aggression, prosocial behaviour and empathy in similar ways, or it could even result from teachers' biases in ratings, favouring girls.
3.1.6 Associations by age

Pupils who were older tended to be rated higher on both aggression and prosocial behaviours at the beginning of the year but not the end of the year, by a small but significant amount.

Older pupils within the intervention group had less of a positive change in affective empathy.

Thus, the intervention may be less effective in increasing older pupils’ affective empathy.

On the post-teacher-report questionnaire, pupils’ date of birth was provided for almost all pupils. A new variable was manually created that calculated number of months old of the pupil, as of June 2015, based on their birth month and year. Therefore, for example, a pupil born Feb 2005 would be 10 years and 4 months, or 124 months. The ages of pupils ranged from 65 months (5 ½ years old) to 126 months (10 ½ years old) with an average age of 94 months (7 years 10 months).

In order to examine the relationship between the empathy, aggression, and prosocial variables and the pupils’ age, Pearson’s correlation coefficient was calculated.

The chart below illustrates the different pupil ages in months. The data was available for 667 pupils and the mean age was 94.53 months (SD = 10.85).

Looking at the average scores at the pre-test and post-test, pupils who were older tended to be rated higher on both aggression and prosocial behaviours at the beginning of the year but not the end of the year, and the relationship was small but significant.

There was no relationship between age and ratings of cognitive and affective empathy. The lack could be because teachers are rating within their classrooms, so they are comparing their pupils to other pupils of the relatively same age.

Examining the change scores, pupils who were older tended to have less of a change in prosocial behaviour across the year ($r = -0.08, p < 0.05$) and less of a change in affective empathy across the year ($r = -0.11, p < 0.01$).
The table below shows the correlations between age and the teacher-report variables.

<table>
<thead>
<tr>
<th>Correlation with Age (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre</strong></td>
</tr>
<tr>
<td>Aggression</td>
</tr>
<tr>
<td>Prosocial</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
</tr>
<tr>
<td>Affective Empathy</td>
</tr>
<tr>
<td><strong>Post</strong></td>
</tr>
<tr>
<td>Aggression</td>
</tr>
<tr>
<td>Prosocial</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
</tr>
<tr>
<td>Affective Empathy</td>
</tr>
<tr>
<td><strong>Change from Pre to Post</strong></td>
</tr>
<tr>
<td>Aggression</td>
</tr>
<tr>
<td>Prosocial</td>
</tr>
<tr>
<td>Cognitive Empathy</td>
</tr>
<tr>
<td>Affective Empathy</td>
</tr>
</tbody>
</table>

Note. All significant correlations were weak
* p < .05, ** p < .01, *** p < .001

Of interest is whether the intervention was more effective for older or younger pupils, or if it was equally effective regardless of age.

As shown in the table below, within the intervention group, age did not correlate with changes in aggression, prosocial behaviour, or cognitive empathy, meaning that the changes (e.g., increases in prosocial) were not associated with the pupil’s age.

However, for affective empathy, it appears that there was a weak, negative correlation with age, r = -.13, p < .05. This suggests that older pupils within the intervention group had less of a change in affective empathy. Thus, the intervention may be less effective in increasing older pupils’ affective empathy.
It is to be noted that correlational findings are generally weak, and therefore caution should be used in making any strong assertions. Overall findings did not suggest that age is a strong determining factor in the effectiveness of Roots of Empathy.
3.1.7 Empathy as a mediator of change in aggression and prosocial behaviour

The impact of the intervention on prosocial behaviour occurred because of an increase in affective empathy, which then predicted an increase in prosocial behaviour.

The intervention directly predicted a decrease in aggression, but this did not occur as a result of the change in affective empathy.

The final analysis step is to test whether empathy was a mediator: in other words, to examine whether the intervention leads to positive outcomes (increases in prosocial behavior and decreases in aggression) via improvements in empathy.

As earlier results showed that intervention vs. control pupils did not change in their cognitive empathy, it is clear the affective empathy is the only possible type of empathy that could have mediated change.

Therefore, a model tested the extent to which being in the intervention (Roots of Empathy) or the control (no intervention) leads to changes in affective empathy, and in turn, whether changes in affective empathy lead to decreases in aggression and/or increases in prosocial behaviour.

In order to test whether there is a fully mediated model or partially mediated model, researchers also examined whether being in the intervention or not directly predicts aggression and prosocial behaviour.

Change scores were calculated by taking the post score minus the pre score. Therefore, positive values represent an increase, while negative values represent a decrease. Following is a path diagram showing the model to be tested for the mediation analysis:
The mediation was tested using path analysis with MPlus v6 software (Muthén & Muthén, 2011). The model was “just-identified”, which means that the number of equations is equal to the number of parameters to be estimated and there are zero degrees of freedom, although just-identified models can still be used to estimate the values of coefficients for the paths.

Examining the standardised beta coefficients, it was clear that change in affective empathy did not significantly predict change in aggression (Beta = .04, p = .13), therefore this path was dropped and the model was re-estimated.

The comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) were used to evaluate the model fit. The recommended cutoff criteria for CFI are > .90, for RMSEA are < .08, and for SRMR are < .08 (Hu & Bentler, 1999).

There are three possibilities for the results:

1. the relationship between Roots of Empathy and the outcomes are not mediated by empathy
2. the relationship between Roots of Empathy and the outcomes are partially mediated by empathy
3. the relationship between Roots of Empathy and the outcomes are fully mediated by empathy (there is no direct relationship between the intervention and the outcomes except through empathy)

Early analysis showed clearly that change in affective empathy did not significantly predict change in aggression (Beta = .04, p = .13, therefore this path was dropped and the model re-estimated.

The revised model had an excellent fit: ($\chi^2 (1, N = 661) = 2.27$, CFI = .99, RMSEA = .04, SRMR = .02. 

Being in the intervention (vs. control) led to an increase in change in affective empathy (Beta = .13, SE = .04, p < .01).

Furthermore, change in affective empathy led to a positive increase in change in prosocial behaviour (Beta = .21, SE = .04, p < .001), while being in the intervention had only a marginally significant positive association with change in prosocial behaviour (Beta = .07, SE = .04, p = .09).

Being in the intervention (vs. control) led to a decrease in change in aggression (Beta = -.19, SE = .04, p < .001). Furthermore, changes in aggression were negatively related to changes in prosocial behaviour (Beta = -.23, SE = .04, p < .001) meaning that as one increased, the other decreased.

These relationships are shown the diagram below – the path diagram of the mediation model results.
As part of the model estimation, researchers indicated that they wanted to calculate indirect effects.

The indirect effect from the intervention to change in prosocial behaviour, via change in affective empathy, was small but significant (Beta = .03, SE = .01, p < .01).

The direct effect from the intervention to change in prosocial behaviour was only marginally significant (Beta = .07, SE = .04, p = .09).

There was no indirect effect from the intervention to change in aggression, since the path from affective empathy to aggression was dropped from the model since it was not significant.

However, there was the direct negative relationship between being in the intervention and change in aggression. Overall, these were very small effects, with $R^2$ values of .05 for change in prosocial behaviour, .02 for change in affective empathy, and .04 for change in aggression.
Summary – empathy as a mediator of change

According to the findings, affective empathy did act as a mediator between the intervention and prosocial behaviour.

The intervention predicted an increase in affective empathy, which predicted a further increase in prosocial behaviour. Theoretically, this makes sense and seems to suggest that empathy needs to be instilled among children in order to encourage prosocial behaviour (arguably ‘empathy in action’) to manifest on children’s actions.

The relationship between Roots of Empathy and prosocial behaviour was fully mediated by affective empathy, meaning that the impact of the intervention on prosocial behaviour occurred because of an increase in affective empathy, which then predicted an increase in prosocial behaviour.

However, the relationship between Roots of Empathy and aggression was a direct one, and was not mediated by empathy, meaning that the intervention directly predicted a decrease in aggression, but this did not occur as a result of the change in affective empathy.

Future research may wish to include a measure of behavioural empathy in the teacher report, although it is likely that this measure would overlap substantially with prosocial behaviour, which is why it was not included in the present evaluation.

Furthermore, if an additional analysis was conducted to examine the subset of pupils who may start out highest in aggression or lowest in empathy, then it is possible that there may be more substantial results for the mediation analysis.


3.2 Quantitative findings from administering KEDS protocol with pupils

To ensure a faithful approach to the administration and analysis of the KEDS data, the research team communicated with Dr Corinne Reid, author of the KEDS tool and Dr Helen Davis, co-author of a KEDS paper, to resolve queries around the scoring protocol for KEDS.

Data checking and cleansing took place before using the systematic scoring protocol to arrive at mean scores for each sub-type of empathy:

- Affective
- Cognitive
- Behavioural

3.2.1 Achieved sample

The research team visited 19 schools from five local authority regions to interview pupils from two classes in each school – intervention and control - totalling 38 classes, with 144 pupils at the pre-test and 112 pupils at the post-test.

One of the 20 schools originally expected to take part in the evaluation, one chose not to engage at all.

Therefore, the research team visited 19 schools from five local authority regions to interview pupils from two classes in each school – intervention and control - totalling 38 classes.

A full set of data was collected from 144 pupils at the pre-test and 112 pupils at the post-test. Pupil interview data was collected for 107 pupils for both pre- and post-test.

Of these 107 pupils, there were 26 pupils in the control group and 81 pupils in the intervention group.
3.2.2 Changes in the three variables (affective empathy, cognitive empathy, behavioural empathy) from pre- to post-, by intervention and control

Interviewing pupils with KEDS tool was time intensive but valuable as it helped to triangulate the data (data sourced via more than one method) and provided an alternative where the pupil could interact with a researcher and speak for him- or herself.

There were two areas where the analysis of the KEDS data provided similar results to the teacher-report:

• The intervention appears to have a slight advantage in terms of increases in affective empathy compared to no change in affective empathy for the control group

• When examining those pupils who start out most in need (bottom 50% for empathy), there is indication that pupils who are lowest in cognitive empathy benefit from the intervention and increase more in cognitive empathy than pupils in the control

In order to analyse this data, pupils were split into control and intervention groups and dependent samples t-tests were used to examine within-pupil changes from pre- to post-test.

Examining the whole sample, researchers then used general linear models with a within-subject factor of change in empathy (pre to post), a between-subject factor (intervention vs. control), and the interaction between the two, to examine whether the change from pre to post in cognitive, affective, and behavioural empathy was similar across the intervention and control groups.

It is important to note that the sample sizes were different, with 25 pupils in the control group who had data at pre and post for KEDS compared to 80 pupils in the intervention who had data at pre and post for KEDS, resulting in uneven group sizes. The small sample size, and uneven grouping, makes it more difficult to find statistically significant differences between groups.

However, it is important to note that the sample size is quite large for such a detailed, interactive, one-to-one measure. The research team decided that the time and effort required for the measure should be focused more intensely on those pupils who were part of the Roots of Empathy intervention in order to better understand their changes during the programme. This was also helpful to triangulate the KEDS data with the teacher-report data for the intervention pupils.
**Affective Empathy**

Affective empathy is the extent to which children feel the same feelings as others, so for example, getting sad when another child is sad.

Pupils in the intervention group significantly increased in their level of affective empathy from pre-to post-test across the school year, $t(80) = -4.54, p < .001$.

Pupils in the control group did not significantly change from pre to post, $t(25) = -1.59, p = .12$.

This is shown in the figure below.

A general linear model was conducted with the within-subjects factor of affective empathy (pre and post for each pupil) and the between-subjects factor of the intervention (intervention vs. control).

Analysis determined that there was not a significant interaction ($F = .33, p = .57$), meaning that the change in affective empathy across the school year, based on KEDS, was similar for both the intervention and control groups.

Given the non-normality of the data, the independent-samples Mann-Whitney U test was also conducted to check whether there was a significant difference between control and intervention groups in the change score for affective empathy.

The Mann-Whitney U test also resulted in non-significance, $p = .74$, meaning it cannot be concluded that the change scores were different across intervention and control.
Affective empathy – examining pupils in need

There is the possibility that the intervention would work even more effectively for pupils who need it the most, which would follow from the patterns found in the teacher-report data.

Therefore, pupils were split into two groups (low, high) based on their initial levels of affective empathy.

Then, dependent-samples t-tests were conducted on pupils in the low group to examine the change from pre to post for the intervention group and the control group.

Pupils in the intervention group significantly increased in affective empathy (increase of 1.7, \( t = -5.11, p < .001 \)) and the control group also significantly increased in affective empathy (increase of 2.0, \( t = -3.34, p < .01 \)).

The pattern follows the general pattern for the full KEDS sample.


Cognitive Empathy

Cognitive empathy is the extent to which children understand why other people feel the way they feel, for example, understanding why another child is laughing or crying.

Pupils in the intervention group significantly decreased in their level of cognitive empathy from pre- to post-test across the school year, $t(80) = 2.13, p < .05$.

Pupils in the control group had a similar pattern, where the level of cognitive empathy decreased from pre to post, but it was not a significant change, $t(25) = 1.08, p = .29$.

A general linear model was conducted with the within-subjects factor of cognitive empathy (2 levels, pre and post for each pupil) and the between-subjects factor of the intervention (2 levels, intervention vs. control).

Analysis determined that there was not a significant interaction ($F = .02, p = .90$), meaning that the average change in cognitive empathy was similar for both the intervention and control groups.

Given the non-normality of the data, the independent-samples Mann-Whitney U test was also conducted to check whether there was a significant difference between control and intervention groups in the change in cognitive empathy.

The test was non-significant, $p = .90$, meaning it cannot be concluded that the change scores were different across intervention and control.

These results may be because cognitive empathy was already quite high for both groups at the beginning of the year (above 4.0 on a scale from 1 to 5).
**Cognitive empathy – examining pupils in need**

There is the possibility that the intervention is more effective for pupils who need it the most, following the patterns found in the teacher-report data.

Therefore, pupils were split into two groups (low, high) based on their initial levels of cognitive empathy.

Dependent-samples t-tests were conducted on pupils in the low group to examine differences in the change from pre to post between the intervention and control groups.

The intervention group significantly increased in cognitive empathy (increase of 1.44, $t = -2.21$, $p < .05$) while the control group did not significantly change ($t = -4.4$, $p = .67$).

This is an opposite pattern from the overall sample and suggests that pupils who are lowest in cognitive empathy benefit from the intervention.
**Behavioural Empathy**

Behavioural empathy is the extent to which children demonstrate empathy through their behaviours, for example, comforting another child who is upset.

Pupils in the intervention group had a significant increase in their level of behavioural empathy from pre- to post-test across the school year, $t(80) = -3.22, p < .01$.

Pupils in the control group had a similar pattern, where behavioural empathy rose, but it was a slightly less strong pattern, $t(25) = -2.61, p < .05$.

A general linear model was conducted with the within-subjects factor of behavioural empathy (pre and post for each pupil) and the between-subjects factor of the intervention (intervention vs. control).

Analysis determined that there was not a significant interaction ($F = .33, p = .57$), meaning that the average change across the school year was similar for both the intervention and control groups.

Given the possible non-normality of the data, the independent-samples Mann-Whitney U test was conducted to check for differences between control and intervention groups in the change in behavioural empathy. This test was also non-significant, $p = .58$, meaning it cannot be concluded that the change scores were different across intervention and control.
**Behavioural empathy – examining pupils in need**

As before, pupils were split into two groups (low, high) based on their initial levels of behavioural empathy.

Dependent-samples t-tests were conducted on the low group to examine differences in the change from pre to post between the intervention and control groups.

The intervention group significantly increased in behavioural empathy (increase of 5.8, $t = -6.28$, $p < .001$) as did the control group (increase of 4.6, $t = -2.71$, $p < .05$).

However, the increase was slightly larger for the intervention group, as shown in the figure below.

![Change in KEDS Behavioural Empathy from Pre to Post for Pupils in Bottom 50%](image)
Summary – changes in the three variables from pre- to post

Within the intervention group, the KEDS data suggests that pupils decreased slightly in cognitive empathy and increased in affective and behavioural empathy.

Within the control group, cognitive and affective empathy did not change across the school year, and there was a slight increase in behavioural empathy.

Therefore, based on KEDS data, the intervention appears to have a slight advantage in terms of increases in affective empathy compared to no change in affective empathy for the control group.

Furthermore, the trends indicate that behavioural empathy may have risen more in the intervention group compared to the control group, although it is likely that the sample sizes were too small to determine that the two trends were significantly different.

When examining those pupils who start out with the most maladaptive profiles (bottom 50% for empathy), there is indication that pupils who are lowest in cognitive empathy benefit from the intervention and increase more in cognitive empathy than pupils in the control.

It is noteworthy that this pattern uncovered by KEDS matches the pattern found in teacher reports of pupils’ cognitive empathy.

Pupils who are lowest in behavioural empathy seem to slightly benefit more from the intervention than those in the control.

There are several issues to consider when interpreting these results. Firstly, both groups seemed to have developed positively, with a slight advantage for the intervention group.

This may be explained because some of these traits, such as behavioural empathy, continue to develop naturally as pupils mature and develop socially and emotionally.

Secondly, the similar results may be due to some confounding that occurred due to the design of the study and larger context of the school. These schools often have multiple interventions occurring at the same time, and pupils in the intervention classrooms may interact quite regularly with pupils from control classes, thus there may be “spillover” effects. This is typical in school-based research.

Thirdly, the KEDS is a new measure and has been validated for use with pupils, however, should be interpreted in light of its ongoing development. Pupils may have some difficulty articulating their responses during the KEDS interview.

Considering that the KEDS data collection was always intended only as a complement to the greater numbers of pupils reported on via the teacher questionnaire, the results are promising. Due to the understandably smaller sample sizes (than for the teacher questionnaire), it is difficult to conclude statistically significant differences. That being said, results based on KEDS suggest either no difference between intervention and control or else a positive difference favouring the effectiveness of the intervention.
3.3 Qualitative findings from class teachers, stakeholders & parents

This section amalgamates the findings from qualitative data from three sources:

1. Free text responses within teacher questionnaire
2. Stakeholder interview
3. Parent focus group or parent interview

Anonymous quotations are used to illustrate themes and attributed as follows:

1. Teacher
2. Stakeholder – [role]
3. Parent

3.3.1 Profile of participants

- 31 teachers provided textual comments at baseline
- 29 teachers provided textual comments at follow-up
- 24 stakeholders were interviewed
- 13 parents attended focus groups or were interviewed

Teachers from 31 classes added textual comments within the teacher questionnaire, broken down as follows, for baseline and follow-up.

<table>
<thead>
<tr>
<th>Local authority</th>
<th>Data type</th>
<th>Role</th>
<th>Numbers</th>
<th>Data type</th>
<th>Role</th>
<th>Numbers</th>
<th>Data type</th>
<th>Role</th>
<th>Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>W Lothian</td>
<td>Baseline</td>
<td>Teacher survey</td>
<td>2</td>
<td>Control classroom teacher</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glasgow</td>
<td>Intervention classroom teacher</td>
<td>5</td>
<td>Follow-up teacher survey</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W Dunbartonshire</td>
<td>Baseline teacher survey</td>
<td>4</td>
<td>Intervention classroom teacher</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S Lanarkshire</td>
<td>Baseline teacher survey</td>
<td>2</td>
<td>Control classroom teacher</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highland</td>
<td>Baseline teacher survey</td>
<td>2</td>
<td>Intervention classroom teacher</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Follow-up teacher survey</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Control classroom teacher</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The research team completed 24 depth interviews with a range of stakeholders with various job roles and received an email response from one further respondent. The team interviewed stakeholders from across Scotland, including those from local authorities (LAs) where the evaluation was taking place as well as contacting people from other LAs not involved in the evaluation to date. Some stakeholders had been involved in the Roots of Empathy programme from the beginning (around four years), with others only more recently involved. The table below shows the range of stakeholders involved.
<table>
<thead>
<tr>
<th>Local authority</th>
<th>Data type</th>
<th>Role</th>
<th>Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>W Lothian</td>
<td></td>
<td>ROE instructor</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td>Glasgow</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ROE instructor</td>
<td>1</td>
</tr>
<tr>
<td>W Dunbartonshire</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Educational staff - head / depute</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ROE instructor</td>
<td>1</td>
</tr>
<tr>
<td>S Lanarkshire</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Educational staff - head / depute</td>
<td>1</td>
</tr>
<tr>
<td>Highland</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ROE instructor</td>
<td>1</td>
</tr>
<tr>
<td>Argyll &amp; Bute</td>
<td>Depth interview</td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Educational staff - head / depute</td>
<td>1</td>
</tr>
<tr>
<td>East Ayrshire</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Educational staff - head / depute</td>
<td>1</td>
</tr>
<tr>
<td>Dumfries &amp; Galloway</td>
<td></td>
<td>Educational staff - head / depute</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ROE instructor</td>
<td>1</td>
</tr>
<tr>
<td>Edinburgh</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Educational staff - head / depute</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Educational staff - head / depute</td>
<td>1</td>
</tr>
<tr>
<td>Fife</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Educational staff - head / depute</td>
<td>1</td>
</tr>
<tr>
<td>Inverclyde</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td>N Lanarkshire</td>
<td></td>
<td>Educational staff - head / depute</td>
<td>1</td>
</tr>
<tr>
<td>Perth &amp; Kinross</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td>S Ayrshire</td>
<td></td>
<td>LA staff - link personnel / development officer / EY service mgr / edu psychologist</td>
<td>1</td>
</tr>
<tr>
<td>N/A</td>
<td></td>
<td>Action for Children staff - children’s services manager, service co-ordinator, KKP</td>
<td>3</td>
</tr>
</tbody>
</table>

24

Thirteen *parents* were consulted: three via an in-depth telephone interview and ten via two focus groups, held at Edinbarnet Primary and St Thomas’ Primary in June.

All were female and varied in age and number of children; one was a parent of a child with disabilities; another had a wider experience of the programme as two of her children had been through the Roots of Empathy programme.
### 3.3.2 Expectations of the programme

**All participants felt they understood the aims and objectives of the Roots of Empathy programme.**

All of the participants interviewed felt they understood the aims and objectives of the Roots of Empathy programme well.

All groups – teachers, stakeholders and parents also identified additional aims, as well as the programme’s stated aims of developing children’s understanding of empathy, reducing aggression and increasing prosocial behaviour:

- Increasing young people’s understanding of emotions – those of others and themselves
- Developing a deep understanding of a baby’s development
- Enhancing future parenting skills
- Creating more empathic communities

“Our expectations are the programme will help the children to develop an understanding of their own emotions / empathy and then transfer understanding into being more compassionate and caring towards one another.” (Teacher)

“…creating the parents of tomorrow…” (Parent)

### 3.3.3 Impact of the programme on pupils’ empathy

**Teachers noted examples of the positive effect of the programme on pupils’ affective empathy and cognitive empathy.**

Some teachers from Roots of Empathy classrooms commented on the positive changes in their pupils, regarding empathy - both affective and cognitive:

“It has made them consider the feelings of others in the classroom” (Teacher)

“Before, I would have had to say, ‘Can you see so and so is upset?’ Now, they can see that for themselves.” (Teacher)

“Many are now able to identify temperaments and how their behaviour impacts on others long term.” (Teacher)

“I have seen a huge change in the empathy shared by the children. They are more aware of each other’s feelings and noticing if someone is upset. They were very nurturing to our baby and they show care for each other.” (Teacher)
3.3.4 Impact of the programme on pupils’ aggression

Teachers and stakeholders noted positive effects of the programme on some pupils’ levels of aggression. They gave examples of the pupils’ more constructive approaches to solving differences since taking part in the programme.

*Teachers* believed the programme had had an effect in reducing pupils’ aggression as pupils adopted constructive means of solving their differences:

“…two children who are coping a lot better with disputes…they tend to walk away from situations now.” (Teacher)

One had noted a particularly positive impact for girls:

“…for some of the girls in particular, that could be a wee bit spoilt, selfish and get involved in disputes…it has made a big difference to them. They’ve taken it on board; they’ve taken a closer look at themselves and they’ve actually been behaving a little bit better with others and it’s changed them.” (Teacher)

Some *stakeholders* agreed that the programme had positively affected aggression:

“Some pupils have improved hugely in their aggression” (Stakeholder)

Regarding their own children’s behaviour, most *parents* did not feel there had been much change in aggression. For some this was because they considered their child to be caring and never showing much aggression beforehand.
3.3.5 Impact of the programme on pupils' prosocial behaviour

Some teachers, stakeholders and parents had noticed that pupils were showing more caring, supportive and sharing behaviour with peers. Parents attributed this to the programme.

A number of teachers said they had noticed a change in the children’s prosocial behaviour since they had run the programme:

“They are much quicker to say sorry and recognise what they have done.” (Teacher)

“I have seen a big improvement in their caring and supporting of each other.” (Teacher)

Among stakeholders, some local authority staff reported their own research with eight schools in their area into the effectiveness of the programme. They noted positive results regarding pupils’ understanding of emotions and improved (less aggressive / more sharing) behaviour, with most teachers claiming the programme either had ‘definitely had a positive impact’ or ‘to an extent had a positive impact’.

Some parents were able to identify improvements in their child’s prosocial behaviour, which they attributed to the Roots of Empathy programme:

“…the importance of kindness and behaviour…they are very, very aware of these things and I think the programme helped 100% to achieve that.” (Parent)

“My little boy was talking about it [programme] a lot - what he thinks about it, what he learned from it and he is more aware of being helpful…so has received a few awards for being helpful in the playground.” (Parent)

Others did not report much change, for some because they considered their child caring already. This perhaps echoes the findings from the teacher-report with those most in need gaining most benefits, compared with those already strong in empathy or prosocial behaviour.
3.3.6 Other impact on pupils

Many participants also talked about other positive impacts of the programme on pupils, in addition to changes in the constructs measured in the teacher report or pupil interview.

**Enjoyment**

Stakeholders, parents and researchers observed the pupils’ enthusiasm for and enjoyment of the programme.

Among stakeholders, there was enthusiasm for the way the pupils had reacted to the programme:

“It definitely has met our expectations - lots of children had a positive reaction to it.” (Stakeholder)

Parents illustrated how much their children had enjoyed the programme, by explaining feelings as the programme ended:

“My daughter was crying the day it finished…her saying '[baby] is not coming in anymore…she really enjoyed it.” (Parent)

Researchers also noted during school visits how enthusiastic pupils were about the programme as they were keen to talk about the visits to the classroom of “their baby”.
Emotional literacy

For teachers and stakeholders, their expectations had been met. They were impressed with pupils’ grasp of emotional concepts and increased, correct use of emotional vocabulary.

Perhaps central to any success of Roots of Empathy in reducing aggression and increasing prosocial behaviour is the improved development of pupils’ emotional literacy.

Among teachers and stakeholders, the clear majority felt the programme had met or exceeded their expectations in this regard, with many witnessing or hearing of pupils displaying high degrees of emotional literacy, attributed to the programme by teachers and head teachers.

Class teachers were perhaps best able to observe such changes taking place in their pupils. They noted that pupils had been able to grasp emotional concepts and use emotional vocabulary correctly, despite some teachers fearing at the outset that this would prove too complicated for young pupils.

“The children have really enjoyed having baby [name] in our class. It has helped them think more about how they are feeling and how they can express those feelings.” (Teacher)

“Many are now able to identify temperaments…” (Teacher)

“The children have truly enjoyed…all of the themes of R of E – and have used the learning sessions to open up discussions about feelings of the children’s own and of the baby” (Teacher)

“They have an understanding of their own learning and development and communicate feelings freely.” (Teacher)

From their own vantage point, stakeholders agreed that the programme had given pupils a greater understanding of and ability to express emotions:

“Everyone in the world feels the same. And that’s a wonderful thing for children to know - you’re not alone with this feeling.” (Stakeholder)
Resilience

Stakeholders and parents gave examples of how the programme had helped some pupils to display resilience in the face of difficulties – death of a father, sibling with leukaemia, hearing impairment and difficult family circumstances.

Among stakeholders, some local authority staff gave specific examples where enhanced emotional literacy had helped a pupil’s resilience in the face of difficult challenges.

One recounted how a pupil’s father had been killed on the way to pick up his son from school. The child’s first lesson back after the incident was the Roots of Empathy lesson on safety and the instructor worried that this would be too traumatic for the child to take part in. However, the child thrived in the lesson and was able successfully to express all of their emotions from the previous weeks in a safe environment. The teacher did not think the child would have been able to get through such a difficult time without the programme.

Another stakeholder had spoken to parents of a pupil whose brother had been diagnosed with leukaemia. The parents reported that the pupil had become far more confident expressing their feelings, including talking about the brother’s illness, since the start of the Roots of Empathy programme.

Parents also gave examples of how the programme had helped their child deal with difficult situations:

- One pupil who was hearing impaired had related strongly to the session covering baby sign language

- One pupil had improved their communication skills, despite a difficult year for the family, which had resulted in them being able to play and be more involved with friends. The child now feels more confident and popular

These examples of pupils’ resilience perhaps point to the particular benefits of the Roots of Empathy programme for those in particular need due to their experiences of emotional or social challenges.

A higher priority for emotional wellbeing

Stakeholders had noticed a shift in pupils’ priorities away from materialism towards emotional wellbeing.

Stakeholders talked of a shift in pupils’ priorities away from materialism towards emotional wellbeing. One example occurred when pupils were asked to say what they hoped for their Roots of Empathy baby:

“They talked about [the baby] being included, as opposed to having a bike or lots of money. It was about being loved and having lots of friends.” (Stakeholder)
Knowledge sharing

Stakeholders and parents reported that pupils enjoyed sharing their knowledge about babies, health and wellbeing from the programme with their families.

Stakeholders and parents both remarked on how much pupils enjoyed sharing their newfound interest and knowledge from the programme with their families. For example, one instructor said that the mum of a pupil in her class had changed the way she put her new-born baby to sleep as a result of what the pupil had told her about the disadvantages of babies sleeping on their fronts.

Parents had noticed that their children’s growing confidence as the child shared their knowledge about babies:

“They were telling me how to do things. My daughter would play with her dolls and show me!” (Parent)

Another child warned its parent about the dangers of drinking fizzy pop in relation to bone development.

3.3.7 Sustained impact on pupils

Stakeholders could not anticipate the likely long-lasting impact of the programme on pupils, though many felt that impact on pupils would be retained.

Stakeholders found it difficult to answer questions about the likely lasting impact of the programme on pupils. Many believed that the full impact on the pupils would only become apparent much later in their lives:

“I think it is too early to tell” (Stakeholder)

“It does what it says on the tin. You do see some change in the kids, not loads but you’ll see it when they’re older.” (Stakeholder)

However, many were optimistic that the programme would have a lasting impact, particularly if the messages of the programme are reinforced in later school years:

“My gut says it will.” (Stakeholder)

“I think it will have wider benefits and [be] something which we will continue to refer to.” (Stakeholder)
3.3.8 Impact on other people

On families of pupils experiencing the programme

Stakeholders noted that the programme had helped improve the home-school relationship. A parent noticed that it was easier to resolve issues of sibling rivalry at home, due to the programme.

A couple of stakeholders provided some accounts of how the programme had helped improve or broker the home-school relationship.

“…just had some conversations with some parents which have been very positive.” (Stakeholder)

“Having spoken to a few parents they feel more involved [with the school (s)] because they have come into assemblies to hear about the programme.” (Stakeholder)

A parent reported that sibling rivalry and squabbling had eased, due to the influence of the programme:

“It’s helped me to make them [children] think about…if they do have a bad behaviour towards each other…It’s easier for me to talk about it because they are more receptive - you know?” (Parent)

On other pupils in the school

Stakeholders and teachers observed a ripple effect of the programme, spreading to other pupils in school – transmitted via the programme noticeboard and via contact in the playground.

Another group positively affected by the programme, identified by stakeholders based within schools, is the other children in the school who have not taken part in the programme. Class teachers identified a ripple effect whereby the changed behaviour of the students in the programme has an influence on the other pupils in the school. Teachers also pointed to the baby’s visit and class notice board as emphasising the messages of empathy to those not receiving the programme:

“I think there have been opportunities for other children in the school to be aware of what’s going on and learn from it to some degree…such as a display in the classroom. The children in the similar age group to the class involved - if those children in that class are having a positive impact on their behaviour - then that’s going to make life better for the other children that they’re mixing within the play group…a ripple effect.” (Stakeholder)

“I think it means more to the school. We’ve got a wee baby just now who’s got two sisters in the school…it involves the whole school.” (Stakeholder)

Whilst undoubtedly beneficial that pupils in the non-intervention group may have also received positive impacts from the intervention, this fact may have had a confounding effect, when comparing intervention and control pupils via the quantitative analysis, resulting in lack of sharp distinction of findings when comparing the two groups.
**On teachers in the school**

Some teachers and head teachers had learnt new ways to deliver the curriculum, through the programme.

Among stakeholders, a number of teachers and head teachers commented on the positive impact the programme had had on them - both personally and professionally. In particular, several felt it had changed, enhanced or improved the way they taught:

“It has showed us teachers how to teach other aspects of the curriculum to children. It’s a hands-on way of making it real to children.” (Stakeholder)

“It has really made me think about other ways to teach the kids that I would never have thought about before.” (Stakeholder)

**On the Roots of Empathy baby**

Stakeholders and parents noted the boost to development for the baby.

A number of stakeholders highlighted the benefits of the programme to the healthy development of the Roots of Empathy babies. Parents of the Roots of Empathy babies noted that their baby had more confidence than their older children did at the same age, which they attributed to the experience of engaging with lots of people. Another parent identified that singing the Roots of Empathy song to the baby helped to calm it down when upset, as during a recent aeroplane flight.
3.3.9 Attributing impact to Roots of Empathy

Stakeholders and teachers were unable to attribute impact directly to the Roots of Empathy programme, mainly due to the presence of other programmes in some schools.

When considering attribution of various impacts to the programme, it should be noted that for some schools the Roots of Empathy programme was one of a number of programmes present, alongside others including as Bounce Back\textsuperscript{14}, Paths\textsuperscript{15}, Resilient Kids\textsuperscript{16} and Incredible Years\textsuperscript{17}.

Even though other programmes may have been delivered to different pupils, various teachers identified that the co-existence of these other programmes made it difficult to establish confidently how much the changes in behaviour they had witnessed were all due to the Roots programme.

“We have put many programmes in place this year to improve resilience, empathy and mental health - so it is hard to be specific. But the children are certainly kinder to each other and the one child lacking empathy has improved noticeably over the year.” (Teacher)

The majority of stakeholders were confident that the Roots of Empathy programme had played a significant part in the behavioural and emotional changes witnessed by teachers and that the programmes complement each other:

“If we look at it holistically, it is a very valuable programme…it fits very well alongside other government initiatives such as the Paths and Incredible Years.” (Stakeholder)

\textsuperscript{14} Bounce Back – An Australian programme that helps children develop a stronger sense of wellbeing.
\textsuperscript{15} Promoting Alternative Thinking Strategies – Programme designed to improve student behaviour and social competence
\textsuperscript{16} Resilient Kids – A mindfulness programme aimed at building students’ focus, balance and self-confidence
\textsuperscript{17} The Incredible Years - programme for reducing children’s aggression and behaviour problems and increasing social competence at home and at school.
3.3.10 Process evaluation – what makes Roots of Empathy work

In addition to assessing impact, the research team took the opportunity to explore aspects of “process evaluation” – delivery of the programme - within the qualitative data collection phase. Researchers asked all qualitative participants about the perceived success factors – strengths – of the programme.

Programme format

For most stakeholders, the structure of the programme worked well, with the baby visit reinforcing the classroom sessions on the same theme.

For the majority of stakeholders, the tight structure and format of the programme was appropriate and worked well:

“It’s really important that the children have the lesson first before the baby comes in. I think the organisation of it is good.” (Stakeholder)

“It is very well organised…the lesson before the baby visit is really important to build up to the baby as the focus.” (Stakeholder)

Generally, stakeholders understood the benefits of the strict programme integrity and structured format, but some saw the accompanying lack of flexibility as a potential drawback:

“It’s tightly restricted in its trademark…so there are only certain ways it can be delivered. You have to accept the integrity of how that’s done. So if you buy into it, you buy it hook line and sinker rather than doing it different ways.” (Stakeholder)
**Baby visits**

All groups of participants saw the baby visits as the defining and popular feature of the programme.

*Teachers* identified the baby visits as a particular highlight of the programme:

“Visits from baby were of most benefit and most popular.” (Teacher)

“Children enjoyed actual visits and most developed a bond with the baby…good to see this rather than watch developmental stages on Smartboard.” (Teacher)

*Stakeholders* also deemed the baby visits the defining aspect of the programme. Several mentioned that the pupils quickly learned to respect and care for the baby’s needs, understanding the boundaries of the green mat!

“I’ve thoroughly enjoyed delivering the programme. I think that having the mum and baby in the classroom is just a fantastic way of the children actually seeing for themselves what they’re learning about.” (Stakeholder)

“All themes have had some impact on the children, but especially those sessions when the baby comes in.” (Stakeholder)

“The adoption of a baby is the best bit … a tiny teacher is a fantastic asset.” (Stakeholder)

*Parents*, too, reported that their children had found the baby visits the most popular and effective parts of the programme; through this medium, the children had learnt a lot without realising it. Their children had particularly enjoyed following the baby’s developmental milestones:

“At our house, it was, ’Wee [baby] won’t be coming in again,’ and they did get really involved.” (Parent)

“She was always telling us, ‘Do you know how much she’s grown and how much she’s weighing and she’s giggling?’” (Parent)

**Individual sessions**

Stakeholders appreciated sessions on crying, emotions and safety; parents commented positively on sessions about milestones and communicating.

*Stakeholders* closely involved in the delivery of the programme were able to talk specifically about particular themes and sessions. They thought that those on ‘crying’, ‘emotions’ and ‘safety’ were most useful or enjoyed. One instructor noted that the pupils particularly benefited from the more interactive sessions:

“I think the children enjoy anything that involves themselves - which would be most of the topics!” (Stakeholder)

*Parents* believed that their children had benefited from sessions on ‘milestones’ and ‘communicating’ – with the latter particularly appreciated by parents of children with a communication or sensory disability.
3.3.11 Process evaluation – suggested changes to programme delivery

Researchers also asked participants for any suggested changes to the delivery of the programme, to enhance or consolidate its benefits.

Programme format

Some stakeholders suggested adapting the programme to make it more Scottish.

A couple of instructors mentioned that the programme could be adapted with a more “local” tone, feeling that the language and books were more suited for a Canadian audience:

“Some of the books are quite Canadian in their language and it might be good to put it into a bit more of a local context.” (Stakeholder)

“It would be helpful to make it a bit more Scottish.” (Stakeholder)

Finding a suitable baby and family

A particular challenge perceived by some stakeholders, was finding a suitable baby and family, due to specific age requirements for the baby and the commitment needed from the family. No solutions were suggested.

One particular challenge perceived by several stakeholders was finding a suitable family and baby to be involved. This was mainly due to the specific age requirements for the baby and the commitment needed from the family:

“It has been difficult sometimes finding a baby of the right age in the right area.” (Stakeholder)

“My colleague helped source the baby herself this year because it’s actually quite difficult to do that, because you need a baby at the right age and a mother willing and able to do it timetable-wise.” (Stakeholder)

Rolling out the programme wider

Some stakeholders and parents suggested extending the programme to wider year groups, including perhaps S1 and S2.

Some stakeholders suggested extending the programme to other year groups and to other schools across Scotland. Some wondered whether extending the programme further to S1 and S2 could be helpful. A variation on this suggestion was to embed the learning further by repeating the programme with the same pupils who have already received it, once they reach the start of secondary school.

Parents also wished that the programme could be delivered to a wider age range and felt that pre-teens or young teenagers could benefit from a focus on emotional learning, such as via the Roots of Empathy programme, at that time.

---

18 This may be a misplaced perception, as Action for Children reports that there is a 95% success rate of finding babies to take part.
Focusing the programme narrower – targeting

Some participants suggested the programme could be targeted at those most in need – whether those with disabilities, complex needs, emotional difficulties or from BME communities.

The accepted format is for the Roots of Empathy programme to be delivered in a whole class environment: an inclusive programme as a result of its universal delivery. Nevertheless, a few participants across all groups felt that the programme could be targeted towards those most in need – whether those with disabilities, complex needs, emotional difficulties or from BME communities.

“The programme has delivered many positive outcomes. There are many needs within my class and it would have possibly have had more of an impact with smaller groups. However, then we would have missed out on the bonding opportunity as a class.” (Teacher)

Parental engagement19

Stakeholders and parents were keen that schools should engage with parents more about the programme, with suggestions including:

- An introductory session about the programme at the start of the school year
- A home-school scrapbook about the programme
- Elements of homework linked to the programme
- Texts from school to notify parents of baby visits
- Invitation to parents to view part of a Roots of Empathy session
- Invitation to parents to join baby celebration, if not currently happening

Stakeholders were strongly in favour of increasing levels of engagement with parents about Roots of Empathy to benefit pupils, their families and the home-school relationship.

This currently typically takes place via a letter sent home about the programme at the beginning of the school year and at the end of the programme an invitation to parents to a “baby celebration” session. Some schools also post information about the programme on an easily accessible display board.

Nonetheless, many stakeholders felt that parents could and should be given more information throughout the life of the programme, for example by providing details of each theme, so that the learning can be continued and consolidated within the home environment:

“It would be quite a good idea to give the parents a bit more information about the themes that we’re doing and roughly when they’re going to be covering that...so they can talk at home. They don’t know much of the detail of what we are actually talking about.” (Stakeholder)

“Some of the parents are not aware of the messages being taught. They need reinforcing at home.” (Stakeholder)

Some schools found it difficult to engage parents but felt it worth the effort even so:

19 Parental engagement is at the discretion of the participating schools. Action for Children and Roots of Empathy work closely with schools to assist them where appropriate with respect to parental engagement
“I think it would be a good idea to get parents more involved but I’m not sure how you would do that … it’s difficult getting parent uptake.” (Stakeholder)

Some practical methods were suggested for increasing parental engagement with the programme:

- Offering information sessions at the start of the school year for parents to attend so they could get a clearer understanding of what the programme is about and what it would entail

- Introducing a scrap book whereby children could note down their feelings and thoughts on each topic so parents stay updated as to what their child is learning

- Elements of homework set which were linked to the programme

Similarly, although some parents were content with not knowing too much about the programme, many parents would like to have been more involved. They observed that their children were too young to communicate fully with parents the details of what was happening in school, making it difficult for parents to build on the learning at home:

“We could…talk to them [at home] coz at the moment when they say stuff [about Roots of Empathy] I don’t understand.” (Parent)

Parents gave similar ideas to the stakeholders for boosting parental engagement, with a few extra ideas:

- School could text parents to let them know the dates of the baby visits, as many schools already communicate with parents this way

- School could invite parents into school to view part of a Roots of Empathy session

- School could invite parents to join the baby celebration, where this did not happen already
3.3.12 Overall assessment of the Roots of Empathy programme

Teachers enjoyed the programme and wished it to continue. Most stakeholders had noted powerful benefits from the programme. Parents liked it because their children enjoyed it.

To sum up, teachers were positive about the programme and hoped it would continue:

“I recommend this for the future.” (Teacher)

“Enjoyed by all…Worth the curricular space…” (Teacher)

The overall views of stakeholders are displayed in the word cloud below:

Parents were complimentary about the programme, describing it as innovative, efficient and special. Moreover, they were happy because their children enjoyed it:

“Excited, coz the kids were happy…” (Parent)
3.4 Limitations

In common with much real world research, this study has a number of limitations, often resulting from events out of the control of the research team.

3.4.1 Sampling

Local authorities, schools and classes were all been selected non-randomly to receive the Roots of Empathy programme.

A sample of local authorities and schools receiving the Roots of Empathy programme were selected non-randomly to take part in the evaluation.

3.4.2 Confounding ‘ripple effect’ of intervention on control pupils

As it was not possible to isolate the intervention pupils from the control pupils for the period of the evaluation, there may have been communication of the messages / learnings between intervention and control pupils, for example in the playground, at home or when socialising out of school. This is a common challenge prevalent in school research.

Schools, understandably, did not seek to insulate control pupils from knowledge of the Roots of Empathy programme. Indeed, many schools shared news about the programme with the whole school community on noticeboards and at parents’ evenings.

The research team had hoped that selecting a control class of a different year group (one lower or one higher) than the intervention class could have avoided some of these potential confounding effects. However, head teachers retained the final say on the selection of the control class and there were some instances where the intervention and control classes in a school were the same year group.
3.4.3 Teacher-report questionnaire

**Date of pre-test questionnaire completion**

The evaluation was commissioned very close to the start of the school year. There followed a set-up period where the research team shared information and requirements with evaluation partners, such as Action for Children, and local authorities. In turn, within local authorities, schools and their individual class teachers were recruited to become intervention or control classes. Then data sharing took place between local authorities and the research team, prior to the baseline data collection.

All the schools involved in the evaluation were free to start their Roots of Empathy programme on a date of their choosing. Their programme start dates fell between mid-September and early November.

At the earliest opportunity, the research team contacted schools, at pre-test, to invite class teachers to answer the baseline questionnaire before any Roots of Empathy activity had taken place in the school.

Where this was not possible teachers were asked to rate their pupils retrospectively, as they had been at baseline, before the start of Roots of Empathy in school. Potentially, these retrospective ratings may not be as accurately recalled as ratings taken contemporaneously would have been.

**Changes in class teacher, with different teacher completing questionnaire post-test**

There were two classes taking part in the evaluation who had a change of teacher part-way through the school year. Therefore, the teacher completing the questionnaire and rating the pupils on empathy, aggression and prosocial behaviour was different from the one who had completed the questionnaire at pre-test. In both cases, head teachers tried to mitigate the effects of this by completing the post-test questionnaire alongside the (new) class teacher. However, the head teacher may still not have been able to provide pupil ratings in a manner consistent with the original class teacher.

In addition, anecdotal evidence provided to the research team, suggested that in one of the cases of changed teacher there had perhaps been a disruptive effect to the pupils involved, which manifested in changes to learning and behaviour. It could be speculated that pupils in this class may not have been able to benefit from the Roots of Empathy programme as much as other pupils whose class teacher remained constant throughout the school year.
3.4.4 Pupil interview

Date of pre-test interview completion
The research team ideally wished to collect all pre-test data before the start of any Roots of Empathy activity within school. In most cases, there was very little or no time to arrange researcher visits to the schools before the start of the programme.

Action for Children reported that it is only at the Theme One post-family visit (session 3 out of 27 over the year) that the pupils’ learning about feelings and perspective-taking may start to be integrated. Therefore, it was decided that pupil interview data would ideally be collected before the Theme One post-family visit (session three).

The majority of pupil interviews were achieved before this point, with just two sets of pupil interviews taking place afterwards (by 3 days and 13 days respectively). In these two instances, it is possible that the early sessions of Roots of Empathy had already taken effect, which could perhaps have resulted in them achieving higher scores than they may have done if they had been interviewed earlier, before the programme had started.

Completing a full KEDS interview with each pupil interviewed
At pre-test, a few pupils were not able to respond to the full set of KEDS questions, because the school day ended mid-interview, because the lunch bell rang, because the pupil became inattentive and bored or because the pupil asked to re-join their class for other reasons.

For the follow-up wave, the research team decided to drop the second part of the cognitive empathy question – the prompt – as it slowed down the interviews, with the related risk of losing the attention of pupils and compromising the quality of data gathered towards the second half of the interview.

At analysis, all pre-test and post-test responses were scored according to a consistent protocol, with the extra ‘prompted’ responses to the cognitive empathy question for the pre-test interviews removed for scoring purposes. However, on reflection, the prompt (received at pre-test, but not post-test) could have had a training effect, resulting in the pupils providing more complete responses for the other questions too – affective empathy, behavioural empathy.

3.4.5 Future potential research

This section shows us that it is important to consider early planning and commissioning of research, to allow greater co-ordination between the different players in the evaluation and avoid as many of these limitations as possible.
4. Conclusions

The evaluation findings have a number of implications for practice regarding the Roots of Empathy intervention. These are our conclusions and recommendations, based on the results of the evaluation:

**Impact on pupils and the case for targeting**

- Pupils receiving the Roots of Empathy programme significantly increased their affective empathy and prosocial behaviour across the school year, compared with control pupils, not receiving the programme. Pupils receiving the programme also decreased in aggression, whilst those in control classes increased in aggression across the school year.

- The Roots of Empathy intervention tends to have its greatest impact for children with the lowest empathy and prosocial skills and who are the most aggressive. The qualitative findings suggested that the programme could bestow increased resilience on pupils who were experiencing personal, familial or health-related challenges.

- Accordingly, when administering Roots of Empathy interventions, the study findings support an approach of targeting classes, schools and local authorities where there is evidence of higher aggression, lower empathy and lower prosocial skills among children (such as a large number of detentions, behavioural problems and incidents, and/or records of emotional and behavioural difficulties, among others).

- Boys tend to benefit more than girls, although girls did display reductions in aggression. The qualitative findings suggest that girls gained other softer outcomes including improved resilience or emotional literacy.

**Mediation and mechanisms of change**

- As the study revealed that prosocial behaviour is mediated through (affective) empathy, there is value in instilling affective empathy through programmes or methods such as Roots of Empathy. Achieving empathy itself is not an end, but a means of achieving further prosocial behaviour.

- The Roots of Empathy programme led directly to a decrease in aggression, unmediated by any changes in empathy. There may be other mediators that were not examined and that could explain why aggression decreased as a result of the intervention.

- Qualitative participants identified success factors of the programme to be its structure and format, the central role of the baby visits to the classroom and the key themes covering crying, emotions, safety, communicating and milestones.
Sustained impact on pupils who took part in the programme

- It is beyond the scope of this evaluation to demonstrate whether benefits gained from the programme will be retained by pupils or whether the benefits may even increase if trajectories of change (in empathy, aggression and prosocial behaviour) continue throughout pupils’ development. Qualitative findings indicated the belief in many teachers, stakeholders and parents that the programme would have long-lasting benefits for the participating pupils, but only a long-term evaluation, following the sample a year or more into the future, could provide the answers.

Significant outcomes

This study found significant results showing that the Roots of Empathy intervention ‘worked’. This is a strong finding, given that it can be difficult to achieve such results in social science research when there are so many different factors taking place in pupils’ lives that could have impacts on the outcomes.

The study made a new contribution to the body of evidence, in showing that empathy is a direct outcome from the Roots of Empathy intervention; and that increased empathy led to increased prosocial behaviour.
5. Appendices

5.1 Research tools

5.1.1 Baseline teacher report questionnaire

To be imported from database, or input by teacher

School name
Primary Class
Number of children in class
Date RoE programme started/s (only if in intervention group)

Please complete the following questions for each child in your class. Please note the order of the lists are randomised for each child.

<table>
<thead>
<tr>
<th>Child forename: ..............</th>
<th>Never true of this child</th>
<th>Always true of this child</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child surname: ..............</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Child gender:   M        F</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This child...

- Hits, shoves, or pushes peers
- Initiates or gets into physical fights with peers
- Threatens to hit or to beat up other children
- Tries to dominate or bully peers
- Says supportive things to peers
- Tries to cheer up peers when they are sad or upset about something
- Is helpful to peers
- Is kind to peers
- Gets upset when another person is acting upset
- Cries or gets upset when seeing another child cry
- Can’t understand why other people get upset
- Rarely understands why other people cry

Repeat for every child in class (imported from above)

Finally, what are your overall expectations of the programme (max 30 words) (intervention teachers only)

Thank and close
5.1.2 Follow-up teacher report questionnaire

Q1a
Have you had your final Roots of Empathy session yet in your class (the post-family visit session of the 9th theme entitled ‘goodbye and good wishes’)?

Yes
No
Don’t Know
Route ‘yes’ to Q1b. Route ‘no’ and ‘don’t know’ to Q1c

Q1b
What was the date of the final Roots of Empathy session in your class?

Enforce date format for response
Don’t know

Q1c
What will be the date of the final Roots of Empathy session in your class?

Enforce date format for response
Don’t know

Q2
Please rate the following statements for each pupil in your class, on a scale of 1-5, where 1 is ‘never true of this child’ and 5 is ‘always true of this child’. If a pupil is no longer in your class, please indicate by ticking the box and you will not have to rate the statements for that pupil. Please note the order of the statements will be randomised for each pupil.

One by one, import list of pupil names and pupil IDs and pupil gender from baseline survey completions – in random order. Cannot be altered by respondent.

Pupil no longer in this class
Provide radio button and route past statements for pupil no longer in class

Hits, shoves, or pushes peers
Initiates or gets into physical fights with peers
Threatens to hit or to beat up other children
Tries to dominate or bully peers
Says supportive things to peers
Tries to cheer up peers when they are sad or upset about something
Is helpful to peers
Is kind to peers
Gets upset when another person is acting upset
Cries or gets upset when seeing another child cry
Can’t understand why other people get upset
Rarely understands why other people cry

1 Never true of this child
2
3
4
5 Always true of this child

Ask Q3 of intervention teachers only
Q3
How would you describe any impact that the Roots of Empathy programme in your school this year has had on the children in your class?

Maximum 50 words for verbatim box. Enforce a response

Ask Q4 of all
Q4a
To what extent has the Roots of Empathy programme in your school this year matched your expectations?
Not matched
Matched
Exceeded

Q4b
Please give further comments to explain your answer above.

Maximum 30 words for verbatim box. Do not enforce any response

Q4c
Please describe any suggested improvements to the Roots of Empathy programme, based on your experience in your school this year.

Maximum 30 words for verbatim box. Do not enforce any response

Thank you for your contribution to the evaluation. Action for Children will be notified that you have completed this survey and will provide the gift voucher to you at your school address.

Capture completion date, for analysis purposes
### 5.1.3 Baseline pupil interview

**Kids Empathy Development Scale (KEDS) Script & Response Sheet (baseline)**  
(to be used in conjunction with KEDS Instruction Manual / briefing document)

<table>
<thead>
<tr>
<th><strong>Researcher Name</strong></th>
<th><strong>Date of Interview DD/MM/YY</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>_____ DD _____ MM _____ YY</td>
</tr>
<tr>
<td><strong>Local authority name</strong></td>
<td><strong>School name</strong></td>
</tr>
<tr>
<td><strong>School ID</strong></td>
<td><strong>Roots of Empathy instructor</strong></td>
</tr>
<tr>
<td><strong>Year of Roots of Empathy delivery</strong></td>
<td><strong>Roots of Empathy prog start date</strong></td>
</tr>
<tr>
<td><strong>Class grade / year group</strong></td>
<td><strong>Class = intervention or control</strong></td>
</tr>
<tr>
<td><strong>Class ID</strong></td>
<td><strong>Number of children in class</strong></td>
</tr>
<tr>
<td><strong>Child’s forename</strong></td>
<td><strong>Child’s surname</strong></td>
</tr>
<tr>
<td><strong>Child’s age</strong></td>
<td><strong>Pupil ID</strong></td>
</tr>
<tr>
<td><strong>Child’s gender</strong></td>
<td><strong>Inputted &amp; Q-C (internal use)</strong></td>
</tr>
</tbody>
</table>

Hello, my name is ……………..

**Show ID badge, if appropriate**

What is your name?

I have come to show you some pictures of children playing and doing other things. I would like to find out what you think about the pictures. It isn’t a quiz, so there aren’t any right or wrong answers – it’s just about what you think. I will scribble down a few words while we talk about each picture. It will take us around 10 to 20 minutes.

[Adult helper name] will stay here with us. If you want to stop before the end, just tell me or [adult helper name] – that will be fine.

Is all that ok? Do you have any questions?

[If needed…] Just so I can get to know you a bit better… (choose one icebreaker)
- Tell me about something you enjoyed doing recently / today
- Tell me something funny about yourself
- What is the best / worst thing about school?
- What is your favourite thing to do at the weekend?

Thank you. We are ready to look at the pictures now.

These pictures will tell a story. I will ask you some questions about these pictures. There is no right or wrong answer (so just try your best).

**Q1 Emotional Perspectives Card (screener)**

**Show showcard 1a (faces, with no words)**

These faces show different feelings.
**Point to each face in turn and ask…**

What feeling (emotion) do you think this face shows?

*If the face is correctly identified, write it in below. If the face is labelled incorrectly, say…*

For this activity, we are going to call this face……

*If any faces incorrectly identified, go back to the start, point to the faces in turn and ask to identify again.*

**Proceed with rest of KEDS if child can identify at least – happy, sad, angry**

**Otherwise, thank and close**

*Write in faces correctly identified by child*
Show each picture in turn, showcards 2-13...
Now I am going to show you some pictures like this. This picture tells a story.

Follow the standard questioning format for each picture:
Six of the scenarios have one blank face - with four questions asked.

Six of the scenarios have two blank faces - with two sets of four questions asked, each set from the perspective of one of the people with a blank face. The two blank faces are to test if the child can take multiple perspectives in complex social situations.

The sequence of the three questions is always the same:
- A (affective)
- C (cognitive)
- B (behavioural)

Do not indicate to the child whether their answer is correct or not.

Q A (affective)
How do you think this [girl / boy / man] feels? Pick one of the faces on the card that best matches how she feels.

See showcard 1a
The child may also verbally label the emotion

Q C (cognitive)
Can you tell me why this [girl / boy / man] feels ……?

One prompt to be used:
(It’s hard, isn’t it.) Can you tell me more about what is happening in this picture?
(It’s hard, isn’t it.) What ideas do you have?

If no response after one prompt – move on

Q B (behavioural)
What would you do if you were that [girl / boy / man]?

If child replies by saying they would feel ……., researcher to prompt:
Yes, but what would you do?
Note down child’s responses in table below – and tick if a prompt was used for that question.

<table>
<thead>
<tr>
<th>Items</th>
<th>Response</th>
<th>Score</th>
<th>Prompt used (tick)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Swings – happy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this boy feels?</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>C – Can you tell me why this boy feels ____?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>B – What would you do if you were that boy? [capture any reason given]</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3 Broken arm – sad</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this girl feels?</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>C – Can you tell me why this girl feels ____?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4 Dark room – afraid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this girl feels?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>C – Can you tell me why this girl feels ____?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5a Toy fight boy – angry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this boy feels?</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>C – Can you tell me why this boy feels ____?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>B – What would you do if you were that boy? [capture any reason given]</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5b Toy fight girl – angry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this girl feels?</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>C – Can you tell me why this girl feels ____?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
### Items

<table>
<thead>
<tr>
<th>Prompt used</th>
<th>Response</th>
<th>Score</th>
</tr>
</thead>
</table>

#### B – What would you do if you were that girl?

<table>
<thead>
<tr>
<th>[capture any reason given]</th>
<th>0 1 2</th>
</tr>
</thead>
</table>

#### 6 Watching TV – relaxed

<table>
<thead>
<tr>
<th>A – How do you think this boy feels?</th>
<th>0 1 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C – Can you tell me why this boy feels _____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B – What would you do if you were that boy? [capture any reason given]</td>
<td>0 1 2</td>
</tr>
</tbody>
</table>

#### 7 Jack-in-the box – surprised

<table>
<thead>
<tr>
<th>A – How do you think this girl feels?</th>
<th>0 1 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C – Can you tell me why this girl feels _____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td>0 1 2</td>
</tr>
</tbody>
</table>

#### 8a Ring-a-rosie outcast – sad

<table>
<thead>
<tr>
<th>A – How do you think this girl feels?</th>
<th>0 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>C – Can you tell me why this girl feels _____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td>0 1 2</td>
</tr>
</tbody>
</table>

#### 8b Ring-a-rosie in – happy

<table>
<thead>
<tr>
<th>A – How do you think this boy feels?</th>
<th>0 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>C – Can you tell me why this boy feels _____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B – What would you do if you were that boy? [capture any reason given]</td>
<td>0 1 2</td>
</tr>
</tbody>
</table>

#### 9a Kick fight victim – afraid

<table>
<thead>
<tr>
<th>A – How do you think this boy feels?</th>
<th>0 1 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C – Can you tell me why this boy feels</td>
<td>0 1 2</td>
</tr>
</tbody>
</table>
B – What would you do if you were that boy? [capture any reason given] 3

9b Kick fight aggressor – angry
A – How do you think this girl feels? 0
C – Can you tell me why this girl feels ___? 0
B – What would you do if you were that girl? [capture any reason given] 0

10 Rocking chair – relaxed
A – How do you think this man feels? 0
C – Can you tell me why this man feels ___? 0
B – What would you do if you were that man? [capture any reason given] 0

Items | Response | Score | Prompt used (tick)
---|---|---|---
11a Parent/child father – angry
A – How do you think this man feels? 0
C – Can you tell me why this man feels ___? 0
B – What would you do if you were that man? [capture any reason given] 0

11b Parent/child child – afraid
A – How do you think this girl feels? 0
C – Can you tell me why this girl feels ___? 0
B – What would you do if you were that girl? [capture any reason given] 0

12 Gift unwrapped – surprised
A – How do you think this girl feels? 0
C – Can you tell me
why this girl feels ____?

<table>
<thead>
<tr>
<th>Prompt used</th>
<th>Response</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td>0 1 2 3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13a Sandcastle victim – sad</th>
</tr>
</thead>
<tbody>
<tr>
<td>A – How do you think this girl feels?</td>
</tr>
<tr>
<td>C – Can you tell me why this girl feels ____?</td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Items</th>
<th>Response</th>
<th>Score</th>
<th>Prompt used (tick)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A – Affective, C – Cognitive, B – Behavioural</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Q14 Do you want to tell me anything else about any of the pictures we looked at just now?

Thank you for talking to me. I hope you enjoyed it. It is time to go back to class now.

Researcher Notes and Observations
5.1.4 Follow-up pupil interview

Kids Empathy Development Scale (KEDS) Script & Response Sheet (follow-up)

<table>
<thead>
<tr>
<th>Researcher Name</th>
<th>Date of Interview DD/MM/YY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Local authority name</strong></td>
<td><strong>School name</strong></td>
</tr>
<tr>
<td><strong>Class grade / year group</strong></td>
<td><strong>Class = intervention or control</strong></td>
</tr>
<tr>
<td><strong>Child’s forename</strong></td>
<td><strong>Child’s surname</strong></td>
</tr>
<tr>
<td><strong>Scanned, input (Qa internal use)</strong></td>
<td><strong>Scored (Qa internal use)</strong></td>
</tr>
</tbody>
</table>

Hello, my name is …………………

*Show ID badge, if appropriate*

What is your name?

I have come to show you some pictures of children playing and doing other things. I would like to find out what you think about the pictures. It isn’t a quiz, so there aren’t any right or wrong answers – it’s just about what you think. I will scribble down a few words while we talk about each picture. It will take us around 10 to 20 minutes.

[Adult helper name] will stay here with us. If you want to stop before the end, just tell me or [adult helper name] – that will be fine.

Is all that ok? Do you have any questions?

Q1 Emotional Perspectives Card (screener)

*Show showcard 1a (faces, with no words)*

These faces show different feelings.

*Point to each face in turn and ask…*

What feeling (emotion) do you think this face shows?

*If the face is correctly identified, confirm and move on. If the face is labelled incorrectly, say…*

For this activity, we are going to call this face……[sad / happy / angry / surprised / afraid / relaxed…]"

*If any faces incorrectly identified, go back to the start, point to the faces in turn and ask to identify again.*

*Proceed with rest of KEDS if child can identify at least – happy, sad, angry*

*Otherwise, thank and close*
Show each picture in turn, showcards 2-13...
Now I am going to show you some pictures like this. These pictures will tell a story. I will ask you some questions about these pictures. There is no right or wrong answer (so just try your best).

Follow the standard questioning format for each picture:
(Six of the scenarios have one blank face - with four questions asked. Six of the scenarios have two blank faces - with two sets of four questions asked, each set from the perspective of one of the people with a blank face. The two blank faces are to test if the child can take multiple perspectives in complex social situations.)

The sequence of the three questions is always the same:
- A (affective)
- C (cognitive)
- B (behavioural)

Do not indicate to the child whether their answer is correct or not.

Q A (affective)
How do you think this [girl / boy / man] feels? Pick one of the faces on the card that best matches how she feels.

See showcard 1a
The child may also verbally label the emotion

Q C (cognitive)
Can you tell me why this [girl / boy / man] feels ……?

One prompt allowed:
(It's hard, isn't it.) Can you tell me more about what is happening in this picture?
(It's hard, isn't it.) What ideas do you have?

If no response after one prompt – move on

Q B (behavioural)
What would you do if you were that [girl / boy / man]?

If child replies by saying they would feel ……, researcher to prompt:
Yes, but what would you do?
<table>
<thead>
<tr>
<th>Questions</th>
<th>Response</th>
<th>Score (QA office use)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2 Swings – happy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this boy feels?</td>
<td></td>
<td>0 1</td>
</tr>
<tr>
<td>C – Can you tell me why this boy feels ____?</td>
<td></td>
<td>0 1 2</td>
</tr>
<tr>
<td>B – What would you do if you were that boy? [capture any reason given]</td>
<td></td>
<td>0 1 2 3</td>
</tr>
<tr>
<td><strong>3 Broken arm – sad</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this girl feels?</td>
<td></td>
<td>0 1</td>
</tr>
<tr>
<td>C – Can you tell me why this girl feels ____?</td>
<td></td>
<td>0 1 2</td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td></td>
<td>0 1 2 3</td>
</tr>
<tr>
<td><strong>4 Dark room – afraid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this girl feels?</td>
<td></td>
<td>0 1 2</td>
</tr>
<tr>
<td>C – Can you tell me why this girl feels ____?</td>
<td></td>
<td>0 1 2</td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td></td>
<td>0 1 2 3</td>
</tr>
<tr>
<td><strong>5a Toy fight boy – angry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this boy feels?</td>
<td></td>
<td>0 1</td>
</tr>
<tr>
<td>C – Can you tell me why this boy feels ____?</td>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>B – What would you do if you were that boy? [capture any reason given]</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td><strong>5b Toy fight girl – angry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this girl feels?</td>
<td>0 1</td>
<td></td>
</tr>
<tr>
<td>C – Can you tell me why this girl feels ____?</td>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td><strong>6 Watching TV – relaxed</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this boy feels?</td>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>C – Can you tell me why this boy feels ____?</td>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>B – What would you do if you were that boy? [capture any reason given]</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td><strong>7 Jack-in-the box – surprised</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A – How do you think this girl feels?</td>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>C – Can you tell me why this girl feels ____?</td>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td><strong>8a Ring-a-robbie outcast – sad</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>How do you think this girl feels?</td>
<td>0 1</td>
</tr>
<tr>
<td>---</td>
<td>----------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>C</td>
<td>Can you tell me why this girl feels ____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B</td>
<td>What would you do if you were that girl? [capture any reason given]</td>
<td>0 1 2 3</td>
</tr>
</tbody>
</table>

**8b Ring-a-rosie in – happy**

<table>
<thead>
<tr>
<th>A</th>
<th>How do you think this boy feels?</th>
<th>0 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Can you tell me why this boy feels ____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B</td>
<td>What would you do if you were that boy? [capture any reason given]</td>
<td>0 1 2 3</td>
</tr>
</tbody>
</table>

**9a Kick fight victim – afraid**

<table>
<thead>
<tr>
<th>A</th>
<th>How do you think this boy feels?</th>
<th>0 1 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Can you tell me why this boy feels ____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B</td>
<td>What would you do if you were that boy? [capture any reason given]</td>
<td>0 1 2 3</td>
</tr>
</tbody>
</table>

**9b Kick fight aggressor – angry**

<table>
<thead>
<tr>
<th>A</th>
<th>How do you think this girl feels?</th>
<th>0 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Can you tell me why this girl feels ____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B</td>
<td>What would you do if you were that girl?</td>
<td>0 1 2 3</td>
</tr>
<tr>
<td>10 Rocking chair – relaxed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>A – How do you think this man feels?</td>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>C – Can you tell me why this man feels ____?</td>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>B – What would you do if you were that man? [capture any reason given]</td>
<td>0 1 2 3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11a Parent/child father – angry</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A – How do you think this man feels?</td>
<td>0 1</td>
</tr>
<tr>
<td>C – Can you tell me why this man feels ____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B – What would you do if you were that man? [capture any reason given]</td>
<td>0 1 2 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11b Parent/child child – afraid</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A – How do you think this girl feels?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>C – Can you tell me why this girl feels ____?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>B – What would you do if you were that girl? [capture any reason given]</td>
<td>0 1 2 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>12 Gift unwrapped – surprised</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A – How do you think this girl feels?</td>
<td>0 1 2</td>
</tr>
<tr>
<td>C – Can you tell me why this</td>
<td>0 1 2</td>
</tr>
<tr>
<td>girl feels ____?</td>
<td>0</td>
</tr>
<tr>
<td>------------------</td>
<td>---</td>
</tr>
<tr>
<td><strong>B</strong> – What would you do if you were that girl? [capture any reason given]</td>
<td>0</td>
</tr>
</tbody>
</table>

**13a Sandcastle victim – sad**

| A – How do you think this girl feels? | 0 | 1 |
| C – Can you tell me why this girl feels ____? | 0 | 1 | 2 |
| **B** – What would you do if you were that girl? [capture any reason given] | 0 | 1 | 2 | 3 |

**13b Sandcastle vandal – happy**

| A – How do you think this boy feels? | 0 | 1 |
| C – Can you tell me why this boy feels ____? | 0 | 1 | 2 |
| **B** – What would you do if you were that boy? [capture any reason given] | 0 | 1 | 2 | 3 |

A – Affective, C – Cognitive, B – Behavioural

Thank you for talking to me. I hope you enjoyed it. It is time to go back to class now.
5.1.5 Pupil interview showcards

SHOWCARD 1a

SHOWCARD 1b

© Project KIDS 2012
SHOWCARD 2

© Project KIDS 2012
SHOWCARD 3

© Project KIDS 2012
SHOWCARD 5

© Project KIDS 2012
© Project KIDS 2012
© Project KIDS 2012
SHOWCARD 8
SHOWCARD 9

© Project KIDS 2012
SHOWCARD 11

© Project KIDS 2012
SHOWCARD 13

© Project KIDS 2012
5.1.6 Stakeholder interview

Roots of Empathy – Stakeholder Interview Guide - 2015

Interviewee: 
Interviewer: 
Local Authority: 
Date: 
Time: 

My name is XXXX and I work for Qa Research – an independent research company. As you may already be aware Qa Research have been commissioned by Action for Children to undertake an evaluation of the Roots of Empathy Programme being delivered across Scotland.

The evaluation has involved a mix of methodologies, looking at gathering baseline and follow-up data from teachers and children receiving the intervention.

As part of the evaluation it was felt useful to gather views of a range of stakeholders including Head teachers, Link people, Instructors and so on.

The aim of the interview today is to gather your views on the programme, any improvements that can be made and in particular the impact the programme is having on the schools, pupils and families.

The interview will last approximately 20-30 minutes. Please be assured that although the information from our discussion may feature in the report we will preserve your anonymity. However, if you refer to your organisation or role within the programme some people may be able to identify you. Because everything you say is important and I won’t be able to write it all down, is it ok if I record the interview?

Any questions before we begin?

Background

Q1. It would be useful if you could tell me a bit about your particular role in the Roots of Empathy Programme. 
   (note to interviewer – the roles will be quite different and varied)

Q2. How long have you been involved in the Roots of Empathy programme? 
   (note to interviewer - for Head teachers you may get two responses how long the school has been involved and how long the Head themselves have been involved)

Q3. What is your understanding of the aims and objectives of the Roots of Empathy programme?

Probe: 
   Overall and for schools or individuals
Progress

Q4a. Do you feel the way the programme is set up and the themes that are covered are suitable?

As a reminder:

<table>
<thead>
<tr>
<th>THEMES</th>
<th>Each theme consists of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Meeting the baby</td>
<td>a <strong>pre-family visit</strong> (when the instructor visits the classroom and introduces a theme)</td>
</tr>
<tr>
<td>2. Crying</td>
<td><strong>the family visit</strong> (when the instructor and parent and baby visit the classroom)</td>
</tr>
<tr>
<td>3. Caring and Planning for Baby</td>
<td><strong>the post family visit</strong> (when the instructor returns to the classroom to discuss what has been learnt)</td>
</tr>
<tr>
<td>4. Emotions</td>
<td></td>
</tr>
<tr>
<td>5. Sleep</td>
<td></td>
</tr>
<tr>
<td>6. Safety</td>
<td></td>
</tr>
<tr>
<td>7. Communicating</td>
<td></td>
</tr>
<tr>
<td>8. Who am I?</td>
<td></td>
</tr>
<tr>
<td>9. Goodbye and Good wishes</td>
<td></td>
</tr>
</tbody>
</table>

Q4b. Are there any changes you would like to see?

Q5. Generally, what would you say has worked well so far?

Probe:
- Structure of the sessions
- Particular sessions
- Baby visits / demonstrations

Why??

Q6. What have been the challenges so far?

Note to researcher – this will depend on the type of stakeholder

Probe:
- Overall delivery of the programme
- Engaging schools / staff
- Recruiting mothers and babies
- Logistics of scheduling the visits
- Lack of resource
- Lack of commitment
- Measuring its success
- Embedding the principles across the schools
- Embedding the principles within the home
Expectations / impact / outcomes

Q7. Has the programme from your personal perspective met your expectations? If not, why not?

Probe – is there anything that surprised you or anything you were expecting from the programme that didn’t happen?

Q8 Overall, do you feel the aims and objectives of the programme have been met / are being met?

(Researcher – be aware that some people answering for their schools, some for their LAs, some for the programme overall)

If so, what evidence do you have / why do you say that?

If not, why not (may be too early to tell)??

Q9a. What impact has the programme had on;
(note to researcher this will depend on the type of stakeholder)

- You / your colleagues
- Your school / schools generally
- Pupils engaged in the programme / other pupils in the school
- Families of those children who have been part of the programme (siblings / parents)

Q9b. Has the programme had or do you expect the programme to have wider or lasting impact / benefit?

Q9c. Has anything else been happening in school at the same time that could account for these impacts?

Q10a. More specifically are you aware / have you noticed any of the following within your school / the schools where the programme is being delivered?

- Better behaviour in the classroom / playground
- Reduction in bullying
- Any other examples of reduced aggression
- Increased awareness of others feelings amongst children involved in the programme
- Increased awareness of own emotions amongst the children involved in the programme
- Increased knowledge of child development amongst children involved in the programme
- Any other examples of increased empathy
- Anything else…
Q10b Do you think there are particular themes / sessions which have a more positive impact than others, if so why?

Q11. Are you / your organisation / school measuring the impact or any outcomes at all? If so, how and what have been the results so far?

Q12. Have you had any feedback / formal or anecdotal from other staff in school / parents / families about the impact the programme is having?

Parent engagement

Q13 Do you think it would be useful to engage parents more in the programme to help continue the learning, reinforce the messages at home?

Final comments

Q14. Do you have any further comments regarding the AfC Roots of Empathy programme?

Probe schools: will you be continuing with the RoE programme in your school?

Q15. To sum up if you could use one word /phrase to describe the AfC programme what would be?

THANK AND CLOSE
5.1.7 Parent focus group discussion guide

Roots of Empathy
Parent Focus Group Guide – 2015

Moderator:
Date:
Time:
School:
Local Authority:
Number of participants:

Each participant to complete a ‘participant profile’ form and hand in by the end of the session

My name is XXXX and I work for Qa Research. We have been commissioned by Action for Children to evaluate the Roots of Empathy programme, which is being delivered or has recently been delivered in your child’s school.

The aim of this focus group is to establish your views of the programme, whether you feel it has benefited your child, changed his / her behaviour and / or impacted more widely on your family. The group will last about an hour. You will all receive a £20 cash incentive as a ‘thank you’ for your time.

I would like to record the group discussion today – unless anyone objects? This is just to help us write up the findings, the recording will not be passed on to anyone else. Your responses will remain anonymous - your name will not be used in any reports that are written.

A few points before I begin:

- There are no right or wrong answers we just want to hear what people think
- Please respect each other’s opinions even if you do not agree with them
- Please do not talk over each other – but feel free to make comments or opinions in response to each other’s points

Any questions before we get started?

Introductions (5 mins)

Each participant to introduce themselves
Awareness of the Roots of Empathy programme (10 mins)

Q1. What do you know about the Roots of Empathy programme / do you know what the Roots of Empathy programme is trying to achieve?

For those that don’t give a brief simple explanation
Roots of Empathy is an evidence-based programme in which children learn how to empathise - with the help of a tiny teacher, a baby. A baby and parent visit a classroom throughout the school year and take part in activities and discussions designed to help pupils understand their own feelings and those of others. The aim of the programme is to reduce problem behaviour in the classroom, including fighting and bullying. Over 500,000 children have taken part in the programme worldwide.

Q2. Who if anyone has informed you about the Roots of Empathy programme?

Probe:
School (Headteacher / class teacher)
RoE instructor
Other professional (e.g. social worker, psychologist)
Your child

Q3. Have you received literature / or other information from school about the Roots of Empathy programme and / or your child’s involvement?

Probe:
When (at the beginning, each term etc..) / how often
In what forms, was the info easy to understand?
Would they have like more info / preferred something else?

Q4. Have you been well enough informed about the programme, or do you think your school or RoE should do more to raise awareness of the Roots of Empathy Programme amongst parents of children involved in the programme? If so, in what ways?

Probe:
Information events
Information sent ahead of each session detailing what the sessions will cover
Noticeboards
Opinions of the Roots of Empathy Programme (15 mins)

Q5. Thinking about what you knew about the Roots of Empathy programme before you came here today, if you had to choose one word / one phrase to describe it, what would it be?

ACTIVITY: Moderator to give out pens / post-its to each participant. Participants to write down one word / phrase on post-it, researcher to gather into positive / negative and discuss reasons.

Q6a. Have you been pleased about anything that was going to be covered in sessions, or about the programme in general?

Q6b. Are you or have you been concerned or worried about anything that was going to be covered in sessions, or about the programme in general?

Q7 Does your child tell you about the programme or about the visits?

Probe;
What are their views - do they like it?
What kinds of things do they say?

Impact of the Roots of Empathy Programme (10 mins)

Q8 Based on what your child may have discussed with you (or enthused about), are there any particular sessions / visits that stand out to you? These can be good or bad.

(a) Initially ask unprompted

HANDOUT:
Using a handout detailing a list of themes to help jog their memories

(b) Then ask prompted

<table>
<thead>
<tr>
<th>THEMES</th>
<th>Each theme consists of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. Meeting the baby</td>
<td>a pre-family visit (when the instructor visits the classroom and introduces a theme)</td>
</tr>
<tr>
<td>11. Crying</td>
<td>the family visit (when the instructor and parent and baby visit the classroom)</td>
</tr>
<tr>
<td>12. Caring and Planning for Baby</td>
<td>the post family visit (when the instructor returns to the classroom to discuss what has been learnt)</td>
</tr>
<tr>
<td>13. Emotions</td>
<td></td>
</tr>
<tr>
<td>14. Sleep</td>
<td></td>
</tr>
<tr>
<td>15. Safety</td>
<td></td>
</tr>
<tr>
<td>16. Communicating</td>
<td></td>
</tr>
<tr>
<td>17. Who am I?</td>
<td></td>
</tr>
<tr>
<td>18. Goodbye and Good wishes</td>
<td></td>
</tr>
</tbody>
</table>
Q9  Has the programme impacted on your child’s behaviour / changed child’s behaviour in any way?

Probe for:
Changed behaviour in general – positive / negative
Showing less aggression at home / around others?
Being kinder to siblings?
Reasoning more?

Q10  Have you seen any other benefits / positive impacts on your child, other family members?

Probe for:
Increased desire to go to school
Noticed child’s friends behaving better
Positive behaviour rubbed off on other siblings
Increased awareness of own feelings and emotions
More aware of others’ feelings
Learnt a bit about child development

The Future (5 mins)

Q11  Are there any changes you would like to see being made / suggestions you would like to give for the development and improvement of the Roots of Empathy programme?

Probe for:
Rolled out more widely across the school / across Scotland
Continued through school years, aimed at older / younger children?

Q12  Do you think it would be useful for you as a parent to have more involvement in the programme to help continue the learning, reinforce the messages at home?

Probe for ideas on how this may work

Final comments (5 mins)

Q13  Finally is there anything else you would like to add about the Roots of Empathy Programme?

Thank and end
5.2 Analysis

5.2.1 Data preparation record

The quantitative data was originally prepared by Qa Research in excel format as four separate files:
- Pre- Teacher Report
- Post- Teacher Report
- Pre- KEDS
- Post- KEDS

The four data files were sent to University of Glasgow and uploaded into SPSS version 22 - a statistical software package.

The variables were manually renamed and assigned an indicator based on when they were collected:
- Wave 1 teacher report pre- baseline data (W1)
- Wave 2 teacher report follow-up data (W2)
- Wave 1 KEDS pre- baseline data (W1K)
- Wave 2 KEDS post- follow-up data (W2K)

For example, “Local Authority” was renamed as W1KLocalAuthor for the wave 1 KEDS data.

In several cases, variables were also recoded to fit the formatting requirements of the software package.

New variables were also manually created: for example, date of birth was translated into a simple categorical variable representing age.

The accuracy of the variables was systematically checked and amended as appropriate. For example, the sum of affective, behavioural, and cognitive empathy scores for the KEDS portion were sometimes incorrect on the original baselines and follow-up data files because “0” was included as a value when the pupil was not a participant or dropped out early, and thus these needed to be recoded as ‘missing value’.

The four files were then merged using the assigned unique identifier. All of the data was at the pupil-level.
5.2.2 Interpreting statistics

P-values

The p-value stands for the probability value, a function of the observed sample results that is used for testing a statistical hypothesis. In social sciences, a p-value of .05 is the standard cut-off for the significance level, thus \( p < .05 \) is considered to suggest a significant statistic. If \( p \) is less than or equal to .05, this means that there is a probability of only 5% that the test incorrectly rejected in the null hypothesis. The lower the p-value, then the more confident there is in the results. The statistical tests take into consideration the sample size and adjust the p-value accordingly.

Typically, p-values are reported to indicate different levels of statistical significance:

- \( \geq 0.05 = \text{non-significant (ns)} \)
- \( p < .05 = \text{significant} \)
- \( p < .01 = \text{highly significant} \)
- \( p < .001 = \text{extremely significant} \)

Correlations

Correlation refers to the relationship between two continuous variables, such as age and number of books read.

A positive correlation indicates that as one variable increases, the other also increases. A negative correlation indicates that as one variable increases, the other decreases. Correlations range between 1.00 to -1.00. A perfect positive correlation would be 1.00, a perfect negative correlation would be -1.00, and no correlation would be .00 (where the relationship between variable A and variable B is completely random). One cannot infer causality from correlations.

A correlation test will provide a p-value to signify whether the correlation is significantly different from 0, meaning that the null hypothesis is rejected that there is no relationship - and in other words, indicating that there is a significant relationship between the two variables. When interpreting the strength of the correlation, many researchers refer to effect size guidelines by Cohen\(^{20}\) (1992): .10 is small/weak, .30 is medium/moderate, and .50 is large/strong.

Independent samples t-tests

The purpose of this statistical test is to look at the average difference on a continuous variable between two categorical groups. A difference between groups is considered significant if the p value is < .05.

Dependent samples t-test

The purpose of this statistical test is to look at the difference in levels between two continuous variables within one sample. This is sometimes also called repeated measures or paired samples. In the dependent samples t-test, the two variables should both be measured for each individual. A result is considered significant if the p value is < .05.

---

ANOVA stands for Analysis of Variance. The purpose of this test is to look at the difference on a continuous variable between three or more groups. The test requires one categorical independent variable with three or more groups and one continuous dependent variable that is normally distributed. ANOVAs can be extended. For example, a 2 x 3 ANOVA would consider two categorical variables at the same time in relation to mean differences in an outcome variable. The ANOVA test produces an F score that determines if the groups, overall, are significantly different. If the groups are different, then a post-hoc pairwise comparison test, such as the Tukey range test, is used to determine which of the pairs are significantly different (e.g., with three groups the post-hoc pairwise comparison would examine A vs. B, B vs. C, and A vs. C).

Path analysis

Path analysis is a method that allows for specifying a theoretical model of the relationships between a set of variables and to assess the goodness-of-fit of the model against the data collected (Byrne, 2013). A theoretical model is proposed, data is collected, and then a series of relationships are estimated based on the model proposed. In order to test whether the estimated model has an acceptable fit, there are several goodness-of-fit indices that can be used that have recommended cut-off criteria:

- Comparative Fit Index (CFI) > .90
- Root Mean Square error of approximation (RMSEA) < .08
- Standardized Root Mean Square Residual (SRMR) < .08
- Tucker Lewis Index (TLI) > .90 (Hu & Bentler, 1999)

As part of the model estimation, standardised beta coefficients can be examined that estimate the extent to which one variable predicts another variable. Standardised coefficients show how many standard deviations an outcome variable (such as prosocial behaviour) change per standard deviation increase in a predictor variable (such as empathy). Similar to correlation coefficients, betas range from -1 (indicating a perfect negative prediction) to +1 (indicating a perfect positive prediction) and 0 meaning there is no association between the variables.

Each coefficient has its own p-value, to determine whether the association is significantly different from zero, again, using the standard cut-off of p < .05.

An advantage of this technique is that it allows for the assessment of direct and indirect effects concurrently. A direct effect would be X significantly predicts Z. An indirect effect would be X significantly predicts Y, and Y significantly predicts Z. Thus, complex models with multiple pathways and relationships between numerous variables can be simultaneously estimated.

In order to determine effect size of all of the predictors on each outcome, researchers use R-squared values ($R^2$), called the coefficient of determination, which can be interpreted as the

---

21 Byrne, B. M. (2013). Structural equation modeling with Mplus: Basic concepts, applications, and programming. Routledge

proportion of variance in the outcome variable explained by the predictor variable(s). This could range from 0% to 100% of variance explained.

Human behaviour is incredibly difficult to predict, is constantly changing, and is multiplicatively determined by background, personality, family life, environmental impacts, and more, and thus it is common to only find small $R^2$ values in social science research projects.

**Normal distribution of data**

Traditional statistical tests are based on the assumption that data is normally distributed.

Throughout the report non-parametric alternative tests have been reported, in addition to the traditional parametric test, when data did not strongly support the assumption of normality.

**References**


Please note, if further information is needed to understand and interpret the statistical procedures used throughout this report, the following guides are recommended:


Web guide on t-tests and ANOVAs: [http://www.statisticssolutions.com/manova-analysis/](http://www.statisticssolutions.com/manova-analysis/)

An overview of statistical tests in SPSS: [http://www.ats.ucla.edu/stat/spss/modules/stats.htm](http://www.ats.ucla.edu/stat/spss/modules/stats.htm)


5.3 Features of well-designed trials

Prior to commissioning this study, Action for Children had commissioned an evidence review from the Social Research Unit at Dartington. It included a table listing important features of well-designed trials.

The research team has adapted the table with a final column to show the features of this current study, compared to the criteria.

Features of well-designed trials, compared with features of this study

<table>
<thead>
<tr>
<th>Features of &quot;good enough&quot; trials</th>
<th>Features of &quot;best&quot; trials</th>
<th>Features of this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignment to the intervention is at a level appropriate to the intervention</td>
<td>There is a minimum of one long term follow-up (at least 12 months following completion of the intervention) on at least one outcome measure indicating whether results are sustained over time</td>
<td>Meets “good enough” criteria</td>
</tr>
<tr>
<td>There is use of measurement instruments that are appropriate for the intervention population of focus and desired outcomes</td>
<td>The evaluation results indicate the extent to which fidelity of implementation affects the impact of the intervention</td>
<td>Meets “good enough” criteria</td>
</tr>
<tr>
<td>Analysis is based on ‘intent to treat’</td>
<td>Where possible or appropriate there is analysis of the impact on sub groups (e.g. do the results hold up for different age groups, boys and girls, ethnic minority groups?)</td>
<td>Meets “best” criteria – analysis of impact on sub groups included</td>
</tr>
<tr>
<td>There are appropriate statistical analyses</td>
<td>There is verification of the theoretical rationale underpinning the intervention, provided by mediator analysis showing that effects are taking place for the reasons expected</td>
<td>Approaching “best” criteria – mediation analysis tests all of the theoretical rationale and verifies (part of) the theoretical rationale underpinning the intervention, or the assumption that empathy would achieve a positive impact on: reduced aggression &amp; increased prosocial behaviour</td>
</tr>
<tr>
<td>Analysis of baseline differences indicates equivalence between intervention and comparison groups</td>
<td></td>
<td>Meets “good enough” criteria. Does not meet “best” criteria – analysis found some non-equivalence between the groups at the pre-test. To account for this, researchers focused on changes from pre to post, rather than comparing at one point in time, since the groups did not start out equal.</td>
</tr>
<tr>
<td>There is a clear statement of the demographic characteristics of the population with whom the intervention was tested</td>
<td></td>
<td>Meets “good enough” criteria. No “best” criteria provided</td>
</tr>
<tr>
<td>There is documentation regarding what participants received in the intervention and counterfactual conditions</td>
<td>Partially meets “good enough” criteria. Researchers not privy to whether participants in the intervention condition received ‘full’ or ‘partial’ Roots of Empathy delivery, i.e. whether due to incomplete delivery or pupil absence. Researchers not privy to whether participants in the intervention and counterfactual conditions also received other relevant interventions via whole-school, whole-class, parental or targeted delivery.</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>There is no evidence of significant differential attrition</td>
<td>Meets “good enough” criteria, with under 5% attrition of the teacher-report data for pupils: 695 at pre, 661 at post. No “best” criteria provided.</td>
<td></td>
</tr>
<tr>
<td>Outcome measures are not dependent on the unique content of the intervention</td>
<td>Meets “good enough” criteria. No “best” criteria provided.</td>
<td></td>
</tr>
<tr>
<td>Outcome measures reflect key developmental outcomes</td>
<td>Meets “good enough” criteria. No “best” criteria provided.</td>
<td></td>
</tr>
<tr>
<td>Measures are not rated solely by the person or people delivering the intervention</td>
<td>Meets “good enough” criteria - different measures and multiple sources of data are considered, participating children and teachers are recipients and observers rather than intervention deliverers. No “best” criteria provided.</td>
<td></td>
</tr>
</tbody>
</table>