education sciences

MDPI

Article

Can We Teach Non-Cognitive Outcomes? A Quasi-Experimental Study of Philosophy for Children

Nadia Siddiqui *, Stephen Gorard and Beng Huat See

Durham University Evidence Centre for Education (DECE), School of Education, Leazes Road, Durham University, Durham DH1 1TA, UK; s.a.c.gorard@durham.ac.uk (S.G.); b.h.see@durham.ac.uk (B.H.S.)
* Correspondence: nadia.siddiqui@durham.ac.uk

Abstract: Non-cognitive learning taking place at school helps form dispositions that can be as important as cognitive outcomes in terms of lifelong relevance. There are diverse interventions and school-based programmes targeting non-cognitive skills, but the evidence of their impact is, so far, unclear. To help increase the evidence bases, we conducted a quasi-experimental study involving 486 pupils in 18 primary schools in North Yorkshire, England, wherein 11 schools participated in Philosophy for Children (P4C), and 7 schools formed a control group that received lessons as normal. At the baseline, the two groups were not equivalent, which means the results need to be treated with appropriate caution. Pupils who received the P4C intervention scored higher on all attitudes and views than they had at the start and improved more than comparator pupils in terms of empathy and fairness. However, they were behind the comparator group in terms of teamwork and democracy, and there was little difference in terms of their self-reported ability to communicate with others. In general, the P4C approach was found to be feasible, and was generally liked by teachers and pupils. Teachers reported improvements in pupil conduct and confidence in P4C sessions and in other learning activities. The indications are that non-cognitive skills are potentially malleable and might be improved through a dialogic approach, such as P4C.

Keywords: non-cognitive outcomes; attitudes; quasi-experiment; philosophy for children

Citation: Siddiqui, N.; Gorard, S.; See, B.H. Can We Teach Non-Cognitive Outcomes? A Quasi-Experimental Study of Philosophy for Children. *Educ. Sci.* **2022**, *12*, 322. https://doi.org/10.3390/educsci12050322

Academic Editor: James Albright

Received: 13 April 2022
Accepted: 29 April 2022
Published: 4 May 2022

1. Introduction

Schools are important for children's academic learning and in helping their interaction with the wider world. In school, children might learn skills to develop trust, judgement, opinions of others, and styles of communication. These early school life experiences might influence the learning process and lifelong attitudes to education. Children spending time with supportive peers in a positive school environment can help them to nurture effective social emotional skills. Some evidence has suggested that children's active participation in school-based activities is inversely associated with risk-taking behaviour [1,2]. Learning about trust, mutual respect, empathy, and fairness happens in schools [3].

Emerging evidence suggests that non-cognitive outcomes, such as enjoyment, happiness, confidence, self-esteem, or determination, can be a foundation of higher attainment [4]. While academic achievements may facilitate access to pathways of success in life, some commentators suggest that there are underlying attitudes and behaviours, also influenced by education, which may be important in their own right [5]. The list of such personal qualities is long and could include social and communication skills, resilience, determination, motivation, confidence, self-esteem, and self-efficacy. Various collective terms are used in the literature, such as non-cognitive outcomes, soft skills, personal characteristics, personality traits, life skills, social and emotional skills, and wider educational outcomes. For consistency and clarity, here we use the term non-cognitive outcomes (of education). This paper highlights some of these wider outcomes of schools, and how they might be improved.

This paper considers the impact of Philosophy for Children (P4C), a popular school-based programme, on pupils' communications skills, sense of judgement, empathy and

sense of democratic values. The paper discusses these non-cognitive outcomes and describes the methods used in our quasi-experimental design, data collection and analysis.

The findings are presented along with the process evaluation and impact outcomes of our evaluation. We consider post-intervention differences between treatment groups, teachers' perceived changes in non-cognitive outcomes, and pupils' perceptions. The paper concludes with implications for future research on assessing non-cognitive outcomes, and suggestions for school policy in implementing evidence-based interventions. The study also reports limitations of the design and analysis.

Can We Directly Address Non-Cognitive Outcomes?

Success in life is strongly associated with academic achievements at school, but there are underlying attitudes and behaviours as important for success in a social world [5,6]. The list of personal attitudes and skills could include social emotional learning, academic buoyancy, compassion, motivation, self-reliance, and self-esteem. In the academic literature, these wider outcomes have been combined under umbrella terms such as non-cognitive outcomes, soft skills, personality traits, and wider educational outcomes [7]. In this paper, we use the term non-cognitive outcomes of education for consistency. Our main purpose in this study is to understand the malleability (or not) of these outcomes through a targeted school-based intervention.

Schools provide opportunities for interaction between children from diverse backgrounds and mixed academic abilities. A broad and diverse school curriculum is useful to make education relevant and effective in diverse settings, but there are still gaps in understanding how to do this [8,9]. Teachers' differential treatment of pupils in the same settings can be detrimental for disadvantaged pupils' learning and wider-outcomes [10]. Children observing and experiencing a variety of models of fairness and justice in school contexts can play an important role in their development. This includes living a school life with opportunities to communicate [11], active learning engagement [12], respect from teachers and peers [13], appropriate feedback with encouragement [14], rewards for a wider set of skills and performance [15], choice and freedom to undertake activities [16], and emotional support [17]. Children's general observations of adults being truthful, trustworthy, honest, and fair can encourage pro-social behaviour [18]. There is emerging evidence showing an association between children's general well-being and their exposure to fair and just treatment observed and experienced in schools [19].

Although the existing evidence on the impact of programmes for children's well-being and general development remains weak, schools regularly implement programmes for character development, higher ambitions, social connectedness, well-being, and happiness. However, these programmes are generally implemented in schools without any robust evidence of impact. This paper examines the use of one possible programme in primary schools.

2. Philosophy for Children and Non-Cognitive Skills

P4C is one of the educational approaches for which there is considerable research material published internationally. The approach started gaining attention in the 1980s, and has generated a large number of discussions, debates, and published articles since [20]. Mathew Lipman developed the original P4C programme 40 years ago. The programme centres on philosophical enquiry and dialogue, the ideas and approaches used in Socratic questioning, Vygotsky's constructivism, and Dewey's ideas of democracy in education. The origin lies in an educational experiment conducted in Rand Primary School of New Jersey, USA [21].

P4C is a stepwise dialogic approach implemented in classrooms. Children and teacher(s) sit in a circle and agree on principles of turn-taking and rules for conduct and communication. The teacher introduces a topic in the form of stimulus (e.g., picture, quotation, news item, short story, or poem) and children are asked to develop questions based on it. The questions might not be factual in nature and could lead to conceptual or reflective discussion. Children are encouraged to use the language of dialogue and

argumentation where they build opinions, ask questions, justify points/choices, exemplify, critique, and elaborate their views, responses, and perceptions. Children then display their proposed questions, and there is a voting process to select the question with the most votes, for subsequent discussion. Children are meant to deploy simple argumentation terms and structures, which they have been taught beforehand, to help make their points. They pass control of speaking to one another. The teacher should only intervene where they observe dialogue not building to a useful discussion. After the discussion is completed, the teacher asks children to give their final thoughts, and say whether those thoughts have changed or not due to the discussion. The session is reviewed by all participants to reflect on the quality of discussion.

The main objectives of the P4C programme are to help pupils to learn critical thinking and effective communication skills. The programme does not have a precise framework or syllabus for implementation. The P4C approach can be implemented for all age groups. The P4C teacher training is based on a large variety of examples, and suggestions for teachers that prepare them for organising useful classroom dialogue, preferably aligned with curriculum based activities [22]. Any age-appropriate materials can be adopted which have pupils' interest or relevance with the taught curriculum. The P4C training providers, Society for the Advancement of Philosophical Enquiry and Reflection in Education. (SAPERE) give access to their repository of resource materials and real examples of P4C sessions, so that trained teachers can build further resources and materials for their sessions.

Prior evaluations have reported a rather mixed impact from P4C as a school-based intervention [23–25]. Leaving aside the evidence on academic outcomes, emerging evidence on P4C is indicating benefits for wider outcomes, such as self-esteem [26], self-confidence, active classroom engagement, enjoyment in learning [27,28], critical thinking, creativity, commitment [29,30], social-communication skills [31], social emotional attitudes [32], and motivation to learn, perhaps especially for disadvantaged children [33]. A quasi-experimental study reported improvements in pupils' social skills, concentration, and their ability to question [34–36]. In summary, P4C is a promising approach, but there is a need for a large-scale impact evaluation of P4C on non-cognitive outcomes.

3. Methods Used in the Study

This study is based on a quasi-experimental design. We assessed the impact of the intervention without randomisation of cases into treatment or control groups, using pre–post intervention comparisons of non-cognitive measures [36–38]. In order to increase rigour in the quasi-experimental design, we introduced a comparator group of schools who had volunteered for P4C but did not receive P4C intervention during the period of this study.

Teachers from 11 schools received P4C training in 2018, led by the expert P4C practitioners and master trainers who provided training and resources. There several methods of promoting dialogic practice in classrooms, but for the purpose of impact evaluation, we adhered to only one practice, namely P4C promoted by the organization, Society for the Advancement of Philosophical Enquiry and Reflection in Education (SAPERE). After the training, teachers implemented P4C in their classes. The training was followed by a support day of practice implemented by the trained teacher, observed by an expert P4C trainer in following the protocol developed by the certified developer. A recommended P4C dosage is one P4C session every week. In the formative assessment findings (Section 4), we have discussed dosage and contents of P4C sessions in more detail.

A further seven schools from the P4C list of schools waiting to receive the training were used as a control group. These seven schools formed a clean quasi-experimental comparator, which we knew had not implemented P4C yet, but which were otherwise equivalent to the new treatment group. This was a very cost-effective and ethical approach, requiring only that we collect post-intervention data from the comparator schools in addition to what had already been planned. All 18 schools were from North Yorkshire, in England, with a reasonable proportion of disadvantaged pupils. There was no school dropout.

There were 375 pupils in years 4 and 5 (initially) in the intervention schools with both pre- and post-intervention results for the non-cognitive survey. An additional 13 pupils completed a pre-intervention form, but then either dropped out or moved school. This is just over 3% attrition. The comparator schools provided a further 111 years 5 and 6 pupil responses to the second post-intervention survey. There was no dropout from this group. The analyses are conducted with a total of 486 pupils across two groups.

The intervention was delivered in the treatment schools for 18 months. The research team observed the intervention being implemented in the treatment schools, while some comparison schools were visited for observations in business-as-usual classrooms.

3.1. Assessing Non-Cognitive Outcomes

Our instrument was a pupil survey, which included 11 scaled attitude items. These items were carefully selected to be representative of behavioural constructs used in standardised assessments, such as self-confidence, resilience, and happiness. The instrument has been used successfully many times [34]. A pilot study was conducted to assess readability for pupils, completion time, and format of these survey items. The response scale for all of the items was discussed with primary school teachers in the study and P4C trained staff members. In the pilot study, we found that pupils' could easily understand the response-scale and showed interest in reporting their attitudes and opinions.

Pupils self-reported their attitude or skills on a scale of 0 to 10 (0 identified as not at all true). The scale of 0 to 10 rating is wide, permitting variation in responses, especially over time. Two of the items are randomly reverse-coded so that the socially desirable response would be 0 rather than 10. This was introduced to keep pupils' attention and encourage them to read and respond to the survey carefully. We have used these pupil survey items in a number of successful evaluation studies for the EEF in England, and others, and are, therefore, reasonably sure that the items are age-appropriate, meaningful, and reliable.

The pupil survey items related to social and communication skills, teamwork, and resilience were pre-selected as being particularly associated with the objectives and activities of P4C. This pre-selection of items was meant to reduce the chance of bias created by cherry-picking by giving post hoc attention to items with positive and favourable results once the results were known. The two pre-selected items were selected by the P4C training experts, and the evaluators accepted their selection. The results for these two selected non-cognitive outcomes formed part of the headline impact findings for the study.

We also included four vignettes/stories, and each had three possible outcomes as response scenarios. Pupils were given the choice to select one of the three given options. The vignettes were about imaginary characters (children), and the response statements were neither right nor wrong, but required a judgement based on issues such as empathy or fairness. These four vignettes were also pre-selected for the headline results, because they had less obviously desirable answers.

3.2. Further Data and Methods of Analysis

The P4C sessions were informally observed in eight schools to collect in-depth information on the implementation quality, challenges, and understanding of the intervention mechanism [37]. We recorded in-depth interviews, session observation details, and feedback from teachers and pupils on their perceptions of P4C sessions, and the impact on their learning and interaction. We observed 16 P4C sessions in schools, and interviewed their 20 teachers to obtain evidence on pupils' development of non-cognitive skills, and teachers' experience of implementing P4C. The delivery of the programme in the schools was observed to assess the dosage and challenges of implementation. We also obtained feedback from 45 pupils on their perceptions of P4C sessions, topics, and quality of discussions. The interviews with teachers and pupils were largely informal and conducted whenever we had the opportunity to talk with them after or before the P4C sessions. The schools agreed to support us in collecting data by conducting interviews and session observations. As illustrated below, the ensuing datasets were analysed by theme.

The results for the 11 scaled items are reported as "effect" sizes at the post-intervention stage, based on a survey conducted after 18 months of the intervention. We report the effect sizes, which are the differences between the mean scores for the two comparison groups, divided by their overall standard deviation.

The four vignettes had categorical responses (one of three choices), and we reported the results as frequencies, and as "effect" sizes in terms of odds ratios [38]. The odds ratios (ad-bc) are calculated by comparing the number of cases selecting one of three options, with the number picking either of the other two. We compared the intervention and control groups for the post-intervention responses, and the pre- and post-intervention percentages, which were available for the intervention group only.

4. Findings from the Formative Assessment of P4C Programme

The school teachers, who were volunteers, generally reported that this programme filled an important gap in their existing curriculum. They found it particularly useful for discussing issues in a safe space, perhaps related to behaviour and disciplinary problems, such as bullying, racism, lying and cheating, equality, and fairness. Teachers also reported that the sessions were helpful for improvement in pupils' critical thinking, listening, and respect for others. P4C sessions usually permitted dialogue between teachers and pupils, with both having equal rights and opportunities to question and argue on chosen questions.

Pupils created P4C questions from a given stimulus, such as a picture, story, or short video. The questions were then displayed by the teachers and opened for a blind voting system so that the most voted question could be discussed in the session. The list of questions developed by students to discuss in P4C sessions included:

- Is it fair to tell a lie sometimes?
- What is fear?
- What is kindness?
- Why do we become angry?
- Is freedom absolute?
- Is it OK to judge people on their looks?
- Are grown-ups always trustworthy?
- Is it possible to always be truthful?

These questions clearly indicate an active thought process and the core ideas have the relevance to pupils' critical thinking skills.

School routines are often somewhat rigid and focused on academic activities. Therefore, teachers sometimes found it challenging to convince their school leaders to create a regular space in the timetable. Teachers frequently mentioned that they face time constraints in conducting P4C regularly when there are so many other activities going on. Discussions on deep concepts and philosophical dispositions are not generally practised through subject content.

4.1. Social and Communication Skills

Several P4C teachers found an improvement in pupils' social and communication skills. The teachers' views about the relationship between P4C and social and communication skills may help in understanding the impact results (below).

In the interviews, teachers reported benefits for children's listening abilities. Many teachers said that it improved pupils' levels of patience to listen to others, wait, and then respond. It is rare that school-based activities have the space for pupils to learn and practice such listening skills which otherwise gets neglected. A teacher reported in the interviews that listening to each other in the sessions has helped pupils to develop arguments. She said that when pupils learn to listen to their peers with care and patience they found more points to build their own response. Other teachers commented:

> "Children given more time to talk without me talking as much, and more time to listen to each other."

> "It is sad that we don't have the time to discuss the quality of our talk and P4C session give us an opportunity to really reflect on what we talk and how we talk."
>
> "There is a lot we learn from our children and with our children when we do P4C."
>
> "I have always felt that communication can change our relationship with children. In P4C I observed that children's communication hugely matters for our understanding of their worldview."

4.2. Respect and Behaviour

Teachers reported that after a few sessions they could see pupils' learned new things about their peers and this resulted in showing respect and compassion. Teachers commented that they often observed how the sessions changed the perception of pupils about each other.

> "Children are learning conflict resolution in these sessions."
>
> "Children seem less likely to squabble at play times. The occasions of telling tales has decreased a lot."
>
> "The children are able to express more clearly."
>
> "We often feel that P4C sessions bring down our guards and biases. We behave more compassionately as a result of these sessions."

Teachers frequently reported that pupils' response to different views and opinions became more respectful as P4C was implemented on a regular basis. A teacher reported these changes in their schools:

> "We have a diverse ethnic and religious community in our schools. Children come from different backgrounds. It was very important for us to have a programme like this where we can discuss such differences in a safe space and environment. I felt that P4C session have helped children understand individual differences and given them skills to respect those differences."
>
> "In our sessions, we have observed children becoming confident in raising questions and challenging opinions. We have focused a lot on respect and turn-taking and we can clearly see that children are becoming more aware of respectful communication."
>
> "Some children have surprised me in the sessions as I never expected that they would speak to group with that level of confidence. I think these sessions are helping children struggling to communicate with confidence."

Effective communication practices, such as turn-taking, listening carefully, waiting for the speaker to finish their point, talking slowly, repeating for clarity, and summarising the talking points are the strategies often discussed in the circle of enquiry. The language for building an argument is taught through practice. The teachers also reported that a general improvement in language and pupils' behaviour was observed in other lessons as well.

Teachers reported a general improvement in pupils' social and communication skills, teamwork, and tolerance. Again, these may help in understanding the impact evaluation results.

4.3. Improvement in Reasoning and Thinking Skills

Many teachers commented that pupils learned to think critically as a result of P4C sessions. One stated that pupils learned to change their perspectives on the presentation of new knowledge and were less hesitant in acknowledging that they changed their opinions. Another teacher reported these sessions were a time for reflection and there were always improvements in the quality of P4C session discussions. Teachers' perceptions of the association between P4C and thinking skills are reflected in their following comments:

> "Children have learned how to develop questions for discussions. These questions are imaginative and full of ideas."

> "I have observed that children are becoming more aware of the language use in argumentation and critical thinking."

P4C sessions allowed teacher and pupil communication to be more like a dialogue where both participated more equally in the discussion. Excitement and enjoyment were clearly observable in the P4C sessions. Pupils often commented on the sessions very positively:

> "I enjoy P4C because it is more like talking. We listen to each other carefully. I like listening other children in my class."

> "I say things in P4C because I know that all are listening to me. I speak without any fear. I don't like when people interrupt me."

> "I enjoy the session because we talk on so many things. We talk about life, nature, animals. I love talking about animals."

> "We share feelings. We share stories. I like when we vote for questions."

Pupils also indicated learning about new ideas and talking points in P4C:

> "I like P4C because you can talk about what is true or false. I can share things in my heart."

> "It is enjoyable because we talk different things. We get new ideas from each other. This is different from other classes. We talk and share our opinions and that is good."

However, some pupils clearly reported the following perceptions of the process:

> "Sometimes I feel bored with things we discuss."

> "Some children talk for a long time. I think they should let others speak."

> "We talk on world problems. I feel sad when we talk about children's problems."

These views suggest that the learning is happening in the sessions and pupils' thinking skills are actively promoted by these conversations. Pupils process ideas through listening, speaking, arguing, and raising questions. Learning through dialogue is a process of development where pupils can sometimes face conflicting perspectives. P4C encourages discussions on challenging issues for which teachers are given useful tips and advice in the training sessions.

The success of P4C sessions depends on teachers' preparation, enthusiasm to conduct the enquiry regularly, willingness to accept challenging arguments from pupils, and being aware of personal bias and readiness to accept others' justifications. These areas develop in teaching practice when teachers have the time for preparation, access to a variety of content, and flexibility in embedding P4C on regular basis.

P4C is not linked with any specific literacy and numeracy targets, which makes it hard for teachers to embed in the regular curriculum. Schools are given the flexibility in how to incorporate P4C in their curriculum. The school leaders we interviewed reported that they integrate P4C in subjects, such as History, English, and Physical Social Health Education (PSHE). However, all of them reported challenges in regular implementation when they are under pressure to meet the attainment targets and Ofsted inspections (the education monitoring organisation in England).

A senior management team leader said that the curriculum time needs to be justified with learning goals in the curriculum, where subject learning is dominant and assessed. It is hard to justify any space for P4C unless the sessions improve subject knowledge and practice. The teachers, on the other hand, felt that P4C is a flexible programme and implementation cannot always be for the purpose of attainment and subject knowledge. We observed that P4C was embedded more rigorously in schools where the senior leaders were supportive of this programme and were actively engaged with teachers in the implementation process.

5. Results from the Impact Evaluation

5.1. The Vignettes

The vignettes were short stories with three response options each, and an open comment box in case pupils wanted to elaborate on their response. Pupils were given instructions to read the story carefully and select one of the three options which they considered a fair response to the question. The response options are neither right nor wrong. They are possible ways of making a judgement and the context of fairness might be subjective. However, our selection of a fair response was based on a consideration of equity [3], consensus among the evaluators, education experts, and teachers who have assisted in the development of these items. Moreover, this method of judging pupils' critical thinking and reasoning ability has been practised widely in research [38].

The first vignette was about a pupil struggling at school. We selected one response as suggesting empathy/generosity (it is fair that the teacher should spend more time helping Jacintha, even if the other pupils have to wait), compared to the other two options (Jacintha should work harder, or Jacintha should be taught in a separate class). At the outset only 48% of the intervention pupils chose the first option, compared to 61% subsequently. This gives an odds ratio of 1.69. However, that is only a before and after pattern.

The comparator group, with no pre-score, had 50% of pupils agreeing with the first option in the post-test (Table 1). This is very similar to the intervention group before the intervention, but less than the intervention group after 18 months of P4C. The difference represents an odds ratio of 1.56, in "favour" of the P4C pupils.

Table 1. Vignette on empathy/generosity: Percentage agreeing with the empathy option.

	Empathy Response	Not Empathy Response
P4C pupils	61	39
Comparison pupils	50	50

A second vignette was on children judging fairness (. . . Sam works harder than Peter in his homework assignment but in the end Peter gets higher marks. Is it fairer to be rewarded for hard work or being clever?) The three options were not right or wrong views, but possible scenarios to judge a fairer outcome in this situation ('1. It is not fair that Sam gets a lower mark when he has put in a lot of effort. 2. Peter must be cleverer than Sam, so it is fair that he gets a higher grade even though he didn't put in as much effort. 3. Sam and Peter should get same grade on this homework'). For the analysis we selected option 1 as 'fairer' and options 2 and 3 as 'not so fair'. The intervention group increased their choice of option 1 marginally over time from 48% to 50%.

At post-test, the intervention group were more likely (50%) to pick this option than the comparator group (39%), with an odds ratio 1.57 (Table 2).

Table 2. Vignette on fairness: Percentage agreeing with the fairer option.

	Fairer Option	Not Fairer Option
P4C	50	49
Comparison	39	60

Another vignette on understanding children's sense of judgement about fairness was included (. . . Jenny's family have a different religion to most people in Britain, and they want Jenny to be taught in a school based on that different religion. This means that she will not go to her local school.) The options given were: 1. 'This is not fair because school is one place where people who are different should be able to work alongside each other'. 2. 'This is fair because people who are different should have the opportunity to attend different schools'. 3. 'I cannot decide if it is fair or unfair that Jenny be taught in a school based on that different religion'. In the analysis, we selected option 1 as a fairer response

and options 2 and 3 as less fair. The intervention group increased their choice of the "fairer" option from 53% to 57% over time.

The intervention group also made this choice more than the comparator group (Table 3), with an odds ratio of 1.27.

Table 3. Vignette on fairness: Percentage agreeing with the fairer option.

	Fairer Option	Not Fairer Option
P4C	57	43
Comparison	51	49

The fourth vignette was on children's understanding and value of democratic participation (. . . Three children are willing to become the leader. What is the fairest way of choosing the group leader?) The percentage of children picking the voting response (all children should vote to select the leader) compared to the other two options (2. a teacher should choose the group leader and 3. a name should be blindly picked), decreased after the P4C intervention from 48% to 35%.

The P4C group picked the democratic option less than the comparator group did (Table 4), with an odds ratio of 0.58. The P4C sessions involve pupils 'voting' to select questions for discussions, so it is possible that specific outcomes influenced their choices here. The ideas behind the other vignettes were not featured specifically in P4C classes.

Table 4. Vignette on democracy: Percentage agreeing with voting for the leader.

	Vote for a Leader	Not Vote for a Leader
P4C pupils	35	65
Comparison pupils	54	46

5.2. The Attitude Items

The intervention group reported a slight improvement in social and communication skills over time from a mean of 6.13 to 6.23. However, there was no substantial difference between the two groups (Table 5).

Table 5. Differences in social and communication skills.

I Am Good at Explaining My Ideas to Other People	Post-Intervention Mean	Standard Deviation	'Effect' Size
P4C	6.23	2.28	-
Comparison	6.26	2.13	-
Total	6.10	2.25	−0.01

A similar pattern is observed for the second headline indicator of impact—the self-reported responses on ability to work with others (Table 6). The intervention group improved slightly over time (from a mean of 7.06 to 7.15), but were behind the comparator group (effect size −0.12).

Table 6. Differences in co-operation and teamwork, all pupils.

I Can Work with Someone Who Has Different Opinions to Me	Post-Intervention Mean	Standard Deviation	'Effect' Size
P4C	7.15	2.47	-
Comparison	7.45	2.26	-
Total	7.21	2.43	-0.12

In fact, the same patterns appeared for all 11 items. The intervention group improved over time, but reported scores that were either the same as or worse than the other group.

6. Discussion

The purpose of this evaluation was to observe the implementation of the P4C programme independently and providing a comprehensive report on the process, as perceived and experienced by teachers and pupils. We created two groups of volunteer schools. Randomisation of schools to groups would have been ideal, but the approach we used was quicker, cheaper, and required no school to change their plans or do extra work.

Pupils' who took part in P4C improved their understanding of fairness and empathy. These two non-cognitive outcomes are important for pupils' critical thinking and adopting a rational response to situations in schools and life in general. The formative evaluation shows that P4C is feasible.

6.1. Limitations of the Study

The findings of this quasi-experimental study contribute important evidence on the feasibility of the P4C programme at school level. The study design used a comparison group of schools, which permitted a larger sample size than otherwise. The quasi-experimental design is weak as schools or pupils were not randomised to groups, and the groups were not exactly comparable at the outset. This means that the trustworthiness of the findings is somewhat lower than they would otherwise be, but higher than a simple comparison or a pre/post only comparison [39].

The survey was designed with care, and piloted for accuracy, reliability, and relevance. There are audit trails for each of the attitude items, which were adapted from several large-scale studies. However, assessing non-cognitive outcomes is intrinsically hard because the response to items will depend on pupils' self-reports. We used two reverse-coded items to make pupils focus and respond more carefully on the scale. The vignette items were introduced in the survey so that pupils can respond without thinking about socially desirable responses.

6.2. Suggestions for Research

We used vignettes as survey items with multiple categorical responses, which have been found promising for interest and high completions rates. Pupils responded to these items with interest and care. This approach can be recommended in addition to or instead of standard psychometric scaling. These vignettes have now been widely used by researchers and, as far we know, over 50,000 pupils in 10 countries around the world have responded to these vignettes [3].

The impact outcomes for the evaluation were generosity, fairness, social responsibility, democratic values, social and communication skills, cooperation, and teamwork, as selected by the programme developers. This quasi-experimental study contributes to the existing evidence on the impact of P4C on non-cognitive skills, and also demonstrates that it may possible to change such non-cognitive skills. Future research can investigate the extent of their malleability through more robust research designs.

Teachers mostly attributed improvement in pupils' communication skills to the P4C programme. Teachers shared their observations of children who had struggled to speak confidently, and P4C sessions gave them space and motivation to develop and practice skills for communication. Some teachers also reported improvement in writing skills when pupils practised constructing arguments in the sessions. These are promising impacts, but based on perceptions, which could be assessed more rigorously in a randomised control trial.

6.3. Implications for Policy and Practice

P4C is seen to work as a whole-class approach, when implemented on a regular basis, where pupils can observe the examples of fairness, honesty, empathy, and good behaviour in the session as well as in the school environment. The desired non-cognitive outcomes,

such as empathy, respectful behaviour, and resilience, would only emerge slowly in pupils' behaviour when there is a regularity of the sessions and all adults involved in the process understand and practice what they learn in the session. For example, to expect children to be polite to each other, teachers and other (older) pupils have to also practise it consistently.

The study findings indicate promise of malleability in non-cognitive outcomes and attitudes, and of improvement through implementing school-based programmes. The causal link still needs to be established between non-cognitive and academic outcomes. There is also a gap in evidence that shows how much improvement is possible in academic attainment just by targeting pupils' non-cognitive outcomes. If interventions for the non-cognitive outcomes are shown to improve academic outcomes, then there is a reasonable argument for integrating them in a national curriculum. School leaders have the flexibility to use pupil premium school funds for implementing evidence-based approaches for providing P4C teacher training and support resources. Prior studies have suggested small benefits for literacy and numeracy, and repeated evaluations show that this dialogic approach, if implemented regularly, does no harm to pupils' achievement of academic outcomes. It may not be the best kind of intervention to improve either academic outcomes or non-cognitive ones the most, but it is one of the few that holds genuine promise of improving both.

Author Contributions: Conceptualization, N.S., S.G. and B.H.S.; Methodology, N.S., S.G. and B.H.S.; Software: SPSS version 21.; Formal analysis, N.S. and S.G.; Investigation, N.S., S.G. and B.H.S.; Writing—original draft preparation, N.S. and S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Peter Sowerby Foundation. The grant number is RF230110.

Institutional Review Board Statement: The ethical review was achieved by Durham University School of Education, Ethics committee (Ethic Committee Name: Ethics Committee, School of Education, Durham University, Approval Code: Evaluation North Yorks P4C Hub Project, Approval Date: 12-08-16).

Informed Consent Statement: All school, teachers, and pupils gave informed consent to be included in this study.

Data Availability Statement: The anonymized data is not for public access. Please contact the authors for partial data access and more details.

Acknowledgments: We thank all the school leaders, teachers and pupils who volunteered to take part in study and consented to use their reported information for the research purpose. We specially thank the leads of SAPERE (Grace Robinson, Steve Williams and Bob House) for the delivery of teacher training and supporting the evaluators in collecting the data. We also thank Jane Yates for helping us in school recruitment for this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Garry, J.P.; Morrissey, S.L. Team sports participation and risk-taking behaviors among a biracial middle school population. *Clin. J. Sport Med.* **2020**, *10*, 185–190. [CrossRef] [PubMed]
2. Hahn, R.A. School Segregation Reduces Life Expectancy in the US Black Population by 9 Years. Health Equity. 2022. Available online: https://www.liebertpub.com/doi/full/10.1089/heq.2021.0121 (accessed on 10 April 2022).
3. Gorard, S.; Smith, E. *Equity in Education: An International Comparison of Pupil Perspectives*; Palgrave Macmillan: London, UK, 2010.
4. Lyubomirsky, S.; King, L.; Diener, E. The benefits of frequent positive affect: Does happiness lead to success? *Psychol. Bull.* **2005**, *131*, 803–855. [CrossRef] [PubMed]
5. Gutman, L.; Schoon, I. The Impact of Non-Cognitive Skills on Outcomes for Young People. Education Endowment Foundation. 2013. Available online: http://educationendowmentfoundation.org.uk/uploads/pdf/Non-cognitive_skills_literature_review.pdf (accessed on 4 October 2021).
6. Brunello, G.; Schlotter, M. Non-Cognitive Skills and Personality Traits: Labour Market Relevance and Their Development in Education & Training Systems. IZA Discussion Paper No. 5743. Available online: https://ssrn.com/abstract=1858066 (accessed on 10 April 2022).
7. Bleazby, J. Fostering moral understanding, moral inquiry & moral habits through philosophy in schools: A Deweyian analysis of Australia's Ethical Understanding curriculum. *J. Curric. Stud.* **2020**, *52*, 84–100.

8. Maylor, U. *Diversity and Citizenship in the Curriculum: Research Review*; Research Report-RR819; Department for Education and Skills: London, UK, 2007.
9. Ajegbo, K.; Kiwan, D.; Sharma, S. *Curriculum Review: Diversity and Citizenship*; Department for Education: London, UK, 2007.
10. Davis, L.; Whiteside, J.L.; Cherng, H.Y.S. One Size Fits All? Gender, Race/Ethnicity, and Happiness in Schools. *Teach. Coll. Rec.* **2021**, *123*, 1–28. [CrossRef]
11. Corden, R. *Literacy and Learning through Talk: Strategies for the Primary Classroom*; McGraw-Hill Education: London, UK, 2000.
12. Wan Yusoff, W.M. The impact of philosophical inquiry method on classroom engagement and reasoning skills of low achievers. *J. Curric. Teach.* **2018**, *7*, 135–146. [CrossRef]
13. Smith, E.; Gorard, S. 'Teachers are kind to those who have good marks': A study of Japanese young people's views of fairness and equity in school. *Comp. J. Comp. Int. Educ.* **2012**, *42*, 27–46. [CrossRef]
14. Thompson, A.M.; Wiedermann, W.; Herman, K.C.; Reinke, W.M. Effect of daily teacher feedback on subsequent motivation and mental health outcomes in fifth grade students: A person-centered analysis. *Prev. Sci.* **2021**, *22*, 775–785. [CrossRef]
15. Harlen, W. *The Role of Teachers in the Assessment of Learning*; Pamphlet Produced by the Assessment Systems for the Future Projects (ASF); Assessment Reform Group: London, UK, 2006; pp. 347–364.
16. Gorard, S. The potential determinants of young people's sense of justice: An international study. *Br. J. Sociol. Educ.* **2011**, *32*, 35–52. [CrossRef]
17. Protheroe, N. Emotional support and student learning. *Principal* **2007**, *86*, 50–54.
18. Paulus, M.; Wörle, M.; Christner, N. The emergence of human altruism: Preschool children develop a norm for empathy-based comforting. *J. Cogn. Dev.* **2020**, *21*, 104–124. [CrossRef]
19. Bell, C. "Maybe if they let us tell the story I wouldn't have gotten suspended": Understanding black students' and parents' perceptions of school discipline. *Child. Youth Serv. Rev.* **2020**, *110*, 104757. [CrossRef]
20. Mannion, J. "It's Absolutely Gripping", The Power of Philosophy for Children. Oracy Cambridge. 2020. Available online: https://oracycambridge.org/absolutely-gripping/ (accessed on 21 January 2021).
21. Lipman, M.; Sharp, A.; Oscanyan, F. *Philosophy in the Classroom*; Temple University Press: Philadelphia, PA, USA, 1980.
22. Zulkifli, H.; Hashim, R. Philosophy for children (P4C) in improving critical thinking in a secondary moral education class. *Int. J. Learn. Teach. Educ. Res.* **2020**, *19*, 29–45. [CrossRef]
23. Trickey, S. Promoting Social and Cognitive Development in Schools. Available online: http://www.ep.liu.se/ecp/021/vol1/026?ecp2107026.pdf (accessed on 10 April 2022).
24. García-Moriyón, F.; Rebollo, I.; Colom, R. Evaluating philosophy for children. *Think. J. Philos. Child.* **2005**, *17*, 14–22. [CrossRef]
25. Tian, S.; Liao, P. Philosophy for Children with Learners of English as a Foreign Language. Journal of Philosophy in Schools, 3. 2016. Available online: http://www.ojs.unisa.edu.au/index.php/jps/article/view/1299 (accessed on 10 April 2022).
26. Sasseville, M. Self-esteem, logical skills and philosophy for children. *Thinking* **1994**, *4*, 30–32. [CrossRef]
27. Swain, J.; Cara, O.; Litster, J. Doing Philosophy in Schools: An Evaluation Report. Institute of Education, University of London. 2013. Available online: https://www.philosophy-foundation.org/asset/download/416 (accessed on 10 April 2022).
28. Topping, K.; Trickey, S. Impact of philosophical enquiry on school students' interactive behaviour. *Think. Ski. Creat.* **2007**, *2*, 73–84. [CrossRef]
29. Shahmohammadi, N. A review of effectiveness of teaching philosophy to children on critical thinking skills of sixth-grade students. *Think. Child.* **2020**, *11*, 95–116.
30. Isiklar, S.; Ozturk, Y.A. The effect of philosophy for children (P4C) curriculum on critical thinking through philosophical inquiry and problem solving skills. *Int. J. Contemp. Educ. Res.* **2022**, *9*, 130–142. [CrossRef]
31. Hedayati, M.; Ghaedi, Y. Effects of the philosophy for children program through the community of inquiry method on the improvement of interpersonal relationship skills in primary school students. *Child. Philos.* **2009**, *5*, 199–217.
32. Colom, R.; Moriyón, F.; Magro, C.; Morilla, E. The long-term impact of philosophy for children: A longitudinal study (preliminary results). *Anal. Teach. Philos. Prax.* **2014**, *35*, 50–56.
33. Creative Teaching and Learning. Closing the Achievement Gap with P4C. 2016. Available online: http://www.jamesnottingham.co.uk/media/2344/closing-the-achievement-gap-with-P4C.pdf (accessed on 10 April 2022).
34. Siddiqui, N.; Gorard, S.; See, B.H. Can programmes like philosophy for children help schools to look beyond academic attainment? *Educ. Rev.* **2019**, *71*, 146–165. [CrossRef]
35. Lord, P.; Dirie, A.; Kettlewell, K.; Styles, B. *Evaluation of Philosophy for Children: An Effectiveness Trial*; EEF: London, UK, 2021.
36. Gorard, S.; Siddiqui, N.; See, B.H. Can 'Philosophy for Children' improve primary school attainment? *J. Philos. Educ.* **2017**, *51*, 5–22. [CrossRef]
37. Siddiqui, N.; Gorard, S.; See, B.H. The importance of process evaluation for randomised control trials in education. *Educ. Res.* **2018**, *60*, 357–370. [CrossRef]
38. Lee, C.J.; Goh, E.C. Using vignettes as a 'safe space'for low-income children to discuss sensitive topics in social work assessment. *Child. Youth Serv. Rev.* **2020**, *111*, 104882. [CrossRef]
39. Gorard, S. *How to Make Sense of Statistics*; SAGE: London, UK, 2021.