How to Evaluate the Effectiveness of a School-Based Intervention

This page intentionally left blank

How to Evaluate the Effectiveness of a School-Based Intervention: Evaluating the Impact of the Philosophy for Children Programme on Students' Skills

BY

OURANIA MARIA VENTISTA

United Kingdom – North America – Japan – India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2021

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-80043-003-7 (Print)
ISBN: 978-1-80043-002-0 (Online)
ISBN: 978-1-80043-004-4 (Epub)

ISOQAR certified Management System, awarded to Emerald for adherence to Environmental standard ISO 14001:2004.

Certificate Number 1985
ISO 14001

INVESTOR IN PEOPLE

To my parents and to Grigorios

This page intentionally left blank

Contents

This page intentionally left blank

List of Tables and Flowcharts

Tables

Flowcharts

This page intentionally left blank

Abbreviations

P4C	Philosophy for Children
RCT(s)	Randomised Controlled Trial(s)
TTCT	Torrance Tests of Creative Thinking
HCTA	Halpern Critical Thinking Assessment
CCTT	Cornell Critical Thinking Test
SD	Standard Deviation
NPD	National Pupil Database
FSM	Free School Meals

This page intentionally left blank

Foreword

Professor Chris Brown, Durham University School of Education

The world we inhabit, COVID-19 aside, has had very mixed fortunes over the last decade. On the one hand, technology has enabled us to achieve things our ancestors would never have believed possible: redefining the way people communicate, collaborate, shop, travel, read, research, watch films, gather information, book holidays, bank and so much more (Greengard, 2015). Digital photography and social media have enabled us to capture, represent and share the world in previously unimaginable ways, while tools such as 3D printers allow us to make real our designs for anything, from sculptures to bridges, at the touch of a button. In terms of the economy, artificial intelligence (AI) and mass automation have been transformative, with much routine and low skilled work now undertaken by robots or algorithms. This use of AI is set to continue, with most commentators agreeing that AI will eventually take over many of the tasks machines can perform equally or better than humans: everything from processing insurance claims to space exploration. And this is problematic: with pre-COVID estimates indicating that, over the next 20 years, some 47% of jobs in the United States and 54% of those in Europe will be lost to machines (Bregman, 2018; Frey & Osbourne, 2013).

But what separates humans from machines is our ability to engage in creative thinking – which along with its alter ego, critical problem solving – is about constructing more or less novel ideas, objects or even worlds (du Sautoy, 2019). It is 'imaginative activity fashioned so as to produce outcomes that are both original and of value' (National Advisory Committee on Creative and Cultural Education (NACCCE), 1999): in other words, activity designed to produce ideas that are not only innovative, but also fit for purpose (Durham Commission on Creativity and Education, 2019). What is novel, fit for purpose and indeed somehow satisfying, will depend on different domains: in what Victorian designer William Morris describes as the lesser arts (such as interior decoration), there can be a strong emphasis on novelty. In fields such as architecture, fitness for purpose is generally likely to receive more attention. In all cases, however, our acts of creation elevate, expand and transform what it means to be human.

But if it's creativity that is keeping us one step ahead of the machines, then education systems globally now need to be focussed on actively supporting future citizens to be able to collaboratively engage in critical and creative thought. In other words, we need education systems to now arm students with the capacities required to explore, experiment, try and re-work, make and re-make, explore and

value difference, overcome obstacles and develop and apply knowledge (Newton & Newton, 2018). But we still need practical suggestions for how to make this happen. Also, to have the confidence that if we are to embrace approaches to creative problem solving, that these will make a concrete and positive difference. With this extraordinary new book, Dr Rania Ventista has managed to meet these twin goals. Taking us on a powerful journey of exploration, not only does Dr Ventista showcase, with some considerable skill, how to evaluate educational interventions effectively; she also illustrates in detail why one intervention in particular – *Philosophy for Children* – matters.

The results are a framework that can and should be used by teachers and school leaders when understanding how to allocate scarce resources. But at the same time, how *Philosophy for Children* can enhance creative problem solving and why educators should be embracing this programme to arm their students with the skills they need to navigate the perils and pitfalls the twenty-first century holds. With this, her first book, Dr Ventista has announced to the world her skills as a scholar and leading thinker in this field. I can't wait for what's to come!

References

Bregman, R. (2018). *Utopia for realists and how we can get there*. London: Bloomsbury.

du Sautoy, M. (2019). *The creativity code: How AI is learning to write, paint and think*. London: 4th estate.

Durham Commission on Creativity and Education. (2019). Durham Commission on Creativity and Education. Retrieved from https://www.artscouncil.org.uk/sites/default/files/download-file/Durham_Commission_on_Creativity_04112019_0.pdf. Accessed on March 5, 2021.

Frey, C., & Osborne, M. (2013). The future of employment: How susceptible are jobs to computerisation. Retrieved from https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

Greengard, S. (2015). *The internet of things*. Cambridge, MA: MIT Press.

National Advisory Committee on Creative and Cultural Education (NACCCE) (1999). *All our futures: Creativity and cultural education*. London: DfEE.

Newton, L., & Newton, D. (2018). *Making purposive thought productive*. Ulm: The International Centre for Innovation in Education.

Chapter 1

Introduction

1.1. The Two Themes

One year ago, I decided to write a book so I could share the findings of an evaluation study I conducted about the impact of the Philosophy for Children (P4C) evaluation. I thought that these findings will be valuable for academics, university students and practitioners who are passionate about P4C. As soon as I started writing this book proposal, I had a phone call with a friend, who is a teacher. When she called me, she had just attended a seminar about teaching methods. The trainer discussed teaching methods and explained how his team developed software for Year 1 students. They implemented it to a single classroom and at the end of that year they found that this implementation led to improved results of the students in Maths and Literacy. Therefore, the trainer recommended to the teacher to buy and implement this software if they teach Year 1 students. The trainer said that the following month they would be implementing this to another Year 1 classroom to see whether the findings of the first study would be replicated.

My response to my friend was that the trainer probably aimed to persuade the teachers who received the series of seminars about the effectiveness of the software in order to sell it or maybe the trainer was simply not aware of how to conduct a proper evaluation. They should have mentioned at least the performance of another group to see the progress of the students who did not have access to this software. Furthermore, it is challenging for anyone to remain independent during the evaluation of a programme when they are also those who developed the programme. To remain independent, it is recommended that the evaluator is not the same person who designed the programme (or software) at the first place.

After that phone call, I thought that if teachers and school leaders develop their knowledge about programme evaluation, they can invest resources (time and sometimes money) on something that it is most likely to be effective. Investing their time in methods and programmes that are more likely to work will increase the likelihood of producing better student outcomes. Therefore, I chose to include in the book proposal the rationale of conducting an evaluation of a school-based programme, which I hope will be helpful to the researchers who aim to conduct an evaluation of any programme and to the practitioners to understand the characteristics of a good programme evaluation.

How to Evaluate the Effectiveness of a School-Based Intervention:
Evaluating the Impact of the Philosophy for Children Programme on Students' Skills, 1–8

doi:10.1108/978-1-80043-002-020211001

Therefore, even though I initially decided to write this book to only share the findings of the P4C evaluation, I decided to also introduce the multidimensional framework as a method to conduct an evaluation of a school programme. The two themes of this book are the evaluation of the P4C programme and framework evaluation of school programmes. Potter (2006a), who discussed the evaluation of social programmes, defined social programmes using the phrase 'any kind of organised endeavour' (p. 410). Similarly, I define the school programmes as any kind of organised endeavour and approach adopted in the schools which aim to lead to specific educational outcomes. In this book, I use the terms 'school programme' and 'school-based intervention' interchangeably. I sometimes use the term 'school programme' to avoid overusing the research terminology 'intervention' and be able to highlight the educational and pedagogical aspects.

This book suggests a multifaceted evaluation. It is multifaceted for two reasons. First, this evaluation investigates multiple sources of evidence, which means that the evidence do not come from a single source. Secondly, the evaluation does not examine the impact of the programme on a specific set of skills, but it examines the programme holistically.

This book suggests that different perspectives, methods and educational outcomes should be considered for an effective programme evaluation. This book also introduces a new method as a part of a multifaceted programme evaluation; an evaluation of the school-based intervention based on accepted pedagogical principles. In this way, the programme evaluation does not strictly consider the perspective of 'what works' and the impact of the programme on a few measurable outcomes, but it also involves the examination of the impact of the programme overall and its link with pedagogy.

For this reason, this book does not take the standpoint of qualitative or quantitative research but combines different research traditions. This book does not focus only on measuring the impact of the programme on specific outcomes, but it also involves an evaluation of the programme based on pedagogical principles. This means that this book does not target only researchers who aim to conduct quantitative research or researchers who prioritise qualitative evidence. This book shows the importance of combining different evidence and perspectives for evaluations. Therefore, it moves away from the tradition of only answering the questions of 'what works'. It aims to conclude whether the programme is worth doing according to pedagogical principles and the expected educational outcomes. The framework suggested in this book combines elements of existing methodological approaches discussed in the evaluation research literature, such as those which aim to generalisation, those giving importance to context, those providing information for the policy-perspective and those which suggest that the evaluator cannot be completely neutral towards a programme (Potter, 2006b).

The book demonstrates how this evaluation framework can be implemented in practice. The evaluation of the P4C programme is an example of a multifaceted evaluation of a school-based programme. For the first time, this book presents the findings of a multidimensional evaluation of the P4C programme. This book offers the fullest and most comprehensive presentation of evidence about the effectiveness of the P4C programme. Even though there are other sources presenting

evidence about the effectiveness of the programme, this book summarises evidence about its impact on all students' skills. It also includes the presentation of the findings of the first large-scale study which examined the impact of the programme on creativity, which was a big literature gap for the P4C literature.

To sum up, this book has two main themes. The first theme refers to a framework of how a school-based intervention can be evaluated in a multifaceted way. The second refers specifically to the evaluation of the P4C programme. I decided to present both themes in the same book because a bidirectional relationship between these two main themes emerges. If a reader is interested in reading how a multidimensional school programme evaluation can be conducted, then the evaluation of P4C will be a useful working example to demonstrate how these theoretical evaluation steps look in practice. If a reader is interested in reading the findings about the P4C evaluation, the theoretical discussion of the steps taken to conduct the evaluation can inform the reader about the robustness of the evidence discussed. It is important for the reader to know that even though many publications which refer to school-based interventions try to promote these interventions or/and are usually written by people who created these interventions or train teachers to implement this intervention, this book evaluates the P4C programme from the standpoint of an independent evaluator. Therefore, an objective (to the extent that this is possible) view about the effectiveness of the programme will be presented.

1.2. Two Premises

Having described the two main topics of this book, the two key premises should be presented. These premises refer to two major needs of education, which should first be accepted by the evaluator before proceeding to implementing this evaluation framework. There is a need for evidence-based education and there is a need for skills-based curricula.

1.2.1. Evidence-based Education

According to an evidence-based education, school policy and practice should be justified based on sound evidence (Coe, 1999). Education in England should be evidence-informed which means that the policy-makers and practitioners should base their decisions on evidence about effectiveness. It is useful to know whether a programme works before time and money are spent on its implementation and therefore it is important to examine and combine the available evidence regarding the programme effectiveness.

Evidence-based education has become one of the widely discussed issues in education communities. In recent years, the Education Endowment Foundation funded trials to produce evidence with main interest to reduce the attainment gap between poor students and their peers (Education Endowment Foundation, 2018). However, it also contributed to an increased interest in educational evidence about effective school interventions since the launch of the Teaching and Learning Toolkit in 2011. This toolkit became very popular among teachers in

England and shed light on effective interventions, their impact and costs, according to its creators (Higgins et al., 2016). The Toolkit 'helped to create a more evidence-led culture in the classroom' (Higgins, 2017). The creation of Research Schools in England, the popularity of events such as ResearchED among teachers and their participation in discussions in social media, such as Twitter, show practitioners' willingness to implement evidence-based practice.

Slavin (2002) discussed that randomised and rigorously matched experiments should be the basis for education practice, in the same way that research supports practice in other disciplines, such as medicine. He also mentioned the replicability of those experiments, which is in fact what makes the research findings trustworthy so that the education policy can be based on them. Particularly, in the age of accountability, implementing effective programmes is judged crucial (Slavin, 2002).

Evidence-based education has been linked to randomised controlled trials (RCTs). Morrison (2001) argued that RCTs and meta-analyses have significant flaws, such as the understatement of value of other data sources, the unrealistic simplification and the neglect of context. The evaluation framework suggested in this book aims to address the need for evidence-based education with a framework which takes into consideration other sources of evidence particularly to evaluate the pedagogical foundation of the programme and how the programme should be implemented.

1.2.2. Skills-based Curricula

The multidimensional evaluation framework does not focus only on attainment. This leads to the second premise. Even though most of the evaluations tend to focus on attainment as the primary outcome to examine the impact of a school-based intervention, the evaluation included in this book examines the impact of the programme on other skills, including thinking skills. This book shifts the tradition of simply focussing on attainment as the important educational outcome and aims to demonstrate that there are also other valuable outcomes in education. Programmes that improve thinking or other skills, such as collaboration, are – at least – equally important with those that improve attainment, they have a value and they should be implemented in schools. Since the evaluation of school-based interventions is suggested to include students' skills, it is important to clarify the importance of developing these skills and explain why curricula should be skills-based.

Traditional education is more focussed on the transmission of existing knowledge which is organised into subjects, whilst progressive education is more focussed on the needs and interests of the learner for what is going to be learnt (Pring, 2007). This debate underlies how education perceives truth and knowledge and therefore what should be taught in schools. Traditional education appears to promote the status quo while progressive education is perhaps more future-orientated. Mitra (2000) said that current knowledge will no longer be valid when the pupils leave school, and the students will have to create new paradigms for new problems in their later life. The term '–twenty-first century learning' refers to education which prepares students for the socioeconomic and political context characterised by

globalisation and everchanging digital technology (Benade, 2014). Even though it could be questioned to what extent education should prioritise the employer's requirements, it is probably acceptable that schooling should be sensitive to these needs. For example, there has been a decrease in the need for activities involving manual labour and an increase in the need for cognitive activities.

Hence, the debate should focus on what type of knowledge should be taught, or what 'twenty-first century learning' should involve? Should schools teach students existing knowledge or develop skills? Hirsch (2011) suggested that core knowledge is required either way and argued in favour of the knowledge taken for granted in classroom and society. Therefore, he suggested a curriculum that builds knowledge grade by grade in specific disciplines, such as Maths and Science. I suggest that recall, rote learning and memorisation should stop being the centre of education. I suggest that schools should not implement knowledge-based curricula with the traditional sense. I associate knowledge-based curricula with hidden curricula. A hidden curriculum might promote particular knowledge, work-related behaviour, such as conforming to authority (LeCompte, 1978) and reproduce the status quo. Learners should also be equipped with the skills to evaluate this knowledge. I argue that in progressive education the freedom that the learner receives is less likely to promote aims of hidden curricula because teaching thinking can either be neutral content or even transformative by promoting independent thinking. Nowadays, access to information is relatively easy and the amount of information available online is enormous and growing. Individuals still need knowledge, but they mainly need the skills to be able to search for information, judge its trustworthiness, and process it in an appropriate manner.

Due to fast changes in society and economy, it has been suggested that the knowledge demands for the twenty-first century are not easily predictable, and thus the education system of each country needs to foster critical thinking and creativity (Berliner, 2011). Furthermore, the era of the twenty-first century has been characterised as a post-truth era. Reznitskaya and Wilkinson (2017) recognised that many people appear to think that there is no objective knowledge. For instance, in history there is the idea of historic scepticism and relativism, where historians cannot agree about what happened and there is no objective history (Blake, 1955). If there is no objective knowledge, then there is no need for knowledge-based curricula. Hence, if relativism is accepted, the demand for teaching thinking skills, such as creativity, critical thinking and problem-solving, is crucial.

This should not mean that knowledge should be completely disregarded. Pring (1980) argued that schooling should develop the mind and he suggested that knowledge is necessary for the development of mind. According to Pring, knowledge should focus both on 'knowing that' (propositional knowledge) and 'knowing how'. The multidimensional evaluation of a programme suggested in this book assesses the effectiveness of the programme to develop students' skills of 'knowing how'. This type of knowledge can still be related to the main subjects, such as Literacy and Maths. Students should be taught how to write, how to read, how to do Maths, how to think.

It is important to develop students' skills, such as critical thinking, collaboration and creativity, because these skills can help them later in life. For example,

developing students' collaboration will be a useful skill for their interpersonal relations in their personal and work life. Critical thinkers can evaluate situations and information and distinguish between fake news, distorted views, propaganda and credible sources. The evaluation of credibility of sources, which is part of the critical thinking skill, can lead to a better evaluation of advertisements and thus facilitate consumers' protection. Collaboration has been associated with critical thinking skills. It has been argued that critical thinking skills improve someone's ability to respond constructively to others during conversations and can lead to respect of others' contribution (Lai, 2011). Therefore, it can be argued that critical thinking can result in the development of someone's collaborative skills. In other cases, it was found that collaborative activities can lead to gains in critical thinking skills (Loes & Pascarella, 2017), whilst these skills are sometimes co-developed by the same interventions. For example, a multidisciplinary project-based intervention led to the development of both critical and collaborative skills (Trisdiono, Siswandari, Suryani, & Joyoatmojo, 2019).

The development of creativity should also be taken seriously, since creativity may contribute to the health of the person, both physically and psychologically, adaptability, self-expression, and problem-solving (Runco, 2004). Creativity can also be linked with elements of good citizenship, such as the pursuit of environmental sustainability (Cheng, 2019). The recent report by the Durham Commission on Creativity and Education (DCCE) drew attention to the importance of creativity and provided evidence about its contribution to the identity of individuals, their sense of community, social mobility and well-being.

To summarise, equipping students with these skills in school and adopting skills-based curricula will support their well-being later in life as citizens of democratic societies and as individuals who work and live in a rapidly changed economy and society. Furthermore, assessments evaluating skills like creativity can be less susceptible to influences from socioeconomic status. For example, the assessment of creativity can be fair way without discriminating particular groups of students. Specifically, the assessment of creativity has been found to compensate ethnic, race and gender differences, which are common in the performance of the students in IQ and attainment tests (Kaufman & Sternberg, 2007). Consequently, assessing thinking skills and prioritising them in education might lead to a fairer education.

1.3. The Outline of this Book

So far, I introduced the two main topics of this book. In the next chapter, the next steps that an evaluator can follow to conduct a programme evaluation are briefly discussed. These steps will be analysed and discussed in-depth in the following chapters.

Chapter 3 introduces the P4C programme, which is the programme that is evaluated in this book. P4C programme is introduced as a dialogic, school-based and skills-oriented intervention which aims to improve thinking skills. This chapter introduces and explains the main characteristics of the programme such as Community of Enquiry and the role of the teacher as a facilitator. Considering that the first step of the evaluation is to establish the reasons why a school-based intervention should be implemented, the reader can find the reasons why there

was a need for an evaluation of the P4C programme. This chapter also includes the pedagogical evaluation of the programme and explains how a programme can be evaluated based on pedagogical principles.

Conducting a systematic literature review is the main theme of Chapter 4. This chapter has two parts. In the first part, the main steps to conduct a systematic literature review are presented. In the second part, a synthesis of the available evidence about the impact of the P4C programme on students' skills is presented.

In Chapter 5, ways to generate evidence for the effectiveness of a programme are discussed. When data about the impact of the programme on some specific skills is not available, the researcher will have to design and conduct an evaluation study to measure the programme effectiveness. This chapter has two parts. In the first part, it discusses how an evaluation study can be designed and conducted. This first part briefly presents the main elements of the method of conducting an evaluation study, such as an experimental or quasi-experimental study. In the second part of the chapter, the reader can find out how these steps were implemented in order to evaluate the impact of the P4C programme on thinking skills. A quasi-experimental study with more than 800 students attending primary schools in England was conducted. The steps of conducting the quasi-experimental study are discussed, such as recruitment and analysis. This part of the chapter is particularly helpful to show how practical challenges (e.g. lack of funding, difficulty in recruitment) can be addressed. Furthermore, this chapter presents a cost-effective way of conducting an evaluation study with a robust design. This can be particularly helpful for postgraduate students and researchers with limited or no funding.

In Chapter 6, the reader can find working definitions for two important thinking skills; critical thinking and creativity. These definitions suggest that these skills can be discipline-independent. According to a multidimensional programme evaluation, the programme evaluation should not focus only on attainment. Thinking skills are important educational outcomes. Therefore, this chapter focuses on students' thinking skills. Even though these skills sound 'abstract', these working definitions are useful because they contribute towards their understanding and measurement. There is still no consensus on a definition for these constructs. The working definitions are useful for evaluation purposes since the evaluator will know what to measure. Furthermore, teachers who understand these constructs better may be able to think of designing activities which develop these skills in the classroom. The relationship between critical thinking and creativity is also discussed.

Since measuring the impact of a school-based intervention on thinking skills can be a necessary part of the multifaceted evaluation of that educational programme, in Chapter 7 the assessment of thinking skills is presented. Some existing assessments of thinking skills are briefly examined. The literature review of the existing assessments aims to assist the researchers who would like to use an existing assessment to measure the impact of a programme on students' thinking skills. Buying an existing assessment may be expensive for practitioners or researchers. Therefore, this chapter also includes some cost-effective solutions. This chapter includes some useful designing steps and tips for researchers who want to design their own thinking assessments.

Chapter 8 presents and discusses the findings of the comparative evaluation study. This study examined the impact of P4C on students' critical thinking and creativity.

Chapter 9 discusses the method of secondary data analysis and how it can be used for evaluation purposes. Specifically, this chapter shows how the secondary data analysis can be used in order to conduct a retrospective quasi-experimental design project with a longitudinal approach. Issues related to this analysis are discussed in this chapter. In the second part of the chapter, the findings of the secondary data analysis for the evaluation of the P4C programme are presented.

In Chapter 10, the decision on whether the programme should be implemented or not is discussed. For most of the programmes, there is evidence in favour and against their implementation. This was also the case for the P4C programme. Some evidence showed the development of students' skills whilst other evidence did not. Having presented the different perspectives and evidence coming from different evaluation methods in the previous chapters, all the available evidence about the P4C programme effectiveness are combined and evaluated in this chapter to reach to a conclusion about the value of the programme.

Chapter 11 negotiates the topic of recommendations for programme implementation only if a school programme is decided to be valuable to be implemented. Having decided that the programme should be implemented, it is important to examine the evidence more closely and see under which conditions this programme seems to be effective and under which conditions it is not. Specific recommendations about the implementation of the P4C programme are made. These specific recommendations are made for anyone interested in implementing P4C in their classroom. Recommendations about the practice of the programme are made for teachers, school inspectors, P4C trainers and educational organisations. The recommendations are based on the evidence overall. This means that the recommendations are made both on the existing and new evidence generated as part of the multifaceted programme evaluation.

The last chapter is a summary of the steps that the evaluators can follow for a multidimensional programme evaluation and what this framework suggests for the evaluations conducted by researchers in the future. Based on the literature review, the process evaluation and the outcomes of the evaluation study conducted for P4C, the implications of how thinking can be taught and assessed in schools are discussed.