Character Education Research: The Good, The Bad and the Ugly

Peter Oldham

Shane McLoughlin

Character Education Research: The Good, The Bad and the Ugly

Peter Oldham

Shane McLoughlin

University of Birmingham

Corresponding author: Peter Oldham, poldham610@gmail.com

ORCiDs: Peter Oldham: (0000-0001-6981-3729)

Shane McLoughlin: (0000-0002-2354-5439)

Introduction Over the past 25 years, Character Education has become a major player in the field of education, developing a rich, interdisciplinary corpus. However, no overview of this rapidly evolving field is available, and several previous methodological concerns about the field, including use of valid measures and use of robust quantitative research protocols, remain addressed. The purpose of this review is to fill this gap in the literature, addressing the following research questions: RQ1: What is the current publishing landscape of Character Education research? RQ2: How methodologically sound is Character Education quantitative research? **Methodology** We followed PRISMA guidelines to conduct a scoping review of research falling under the Character Education 'brand'. We searched the Scopus and Web of Science databases, and, following abstract screening, generated a pool of 981 articles for analysis. **Results** RQ1: Character Education research is increasing exponentially, with 1929 authors from 803 institutions in 54 countries publishing in 395 academic journals, supported by 153 funding organisations. However, 38% of the research was of questionable quality. RQ2: 72% of quantitative Character Education studies did not use a valid measure, and, for intervention studies, only 7% used an active control, 70% did not randomise and 86% did not report a standard error, effect size or both. **Recommendations** A greater depth of Character Education research in most contexts, particularly via intervention studies; development of more valid Character-related measures; and intervention studies to increase use of active controls and randomisation, and ensure they report effect size and standard error.

Character Education Research: A Scoping Review

Over the past 25 years, Character Education has taken the world of education by storm. It has become a key component of educational curricula globally, from the UK (Ofsted, 2019) to Indonesia (Gunawan, 2017) to Singapore (Ministry of Education, 2020) to Canada (OCDSB, 2024), and has even developed its own associations[1], specialist research centres[2], conferences [3], and academic journal[4]. Character Education is therefore now a major global educational brand.

Broadly speaking, Character Education can be defined as attempts to teach positive personal characteristics, some of which are imbued with a moral purpose (Kristjánsson, 2015). Such characteristics are multifariously described as traits, strengths, virtues, or values. Due to its dualistic nature, including both positive development and moral elements, Character Education also reaches into the fields of Positive Psychology, Moral Education, and Socioemotional Learning. As well as appearing alongside these closely-related subject areas, Character Education has also been studied alongside a multitude of other educational topics including sustainability (Jordan, 2022), citizenship (Davies et al., 2005), Religious Education (Arthur et al., 2019), educational policy (Arthur, 2005), and Physical Education (Brunsdon, 2023). The novel, inclusive, and diverse nature of Character Education's academic corpus makes it an enticing subject of study for researchers.

---

[1] Association of Character Education*: https://character-education.org.uk/* (Accessed 21/11/24)

[2] Jubilee Centre for Character and Virtues*: https://www.jubileecentre.ac.uk/* (Accessed 21/11/24)

[3] ECVA II Conference - Human Flourishing and Character Education: *https://www.uer.it/en/european-character-and-virtue-association-2024/* (Accessed 21/11/24)

[4] Journal of Character Education: *https://www.infoagepub.com/journal-of-character-education.html* (Accessed 21/11/24)

Previous Character Education Reviews

A number of previous reviews of Character Education research have been undertaken. These studies fall into two broad groups: bibliometric analyses, and meta-analyses.

Three Character Education bibliometric analyses have been published in recent years. Firstly, Sumiharsono et al. (2023) conducted a bibliometric analysis of Character Education research published between 2018 and 2023 from the Scopus database. They investigated a number of key variables including rate of publication, citations, authorship, and institutional affiliations. Their results included that: the rate of publication of Character Education articles appeared to be slowing; the average number of article citations was 3.9; the most published author was Marini, and the most published university was the Universitas Negeri Malang. However, the rigour in this review was unclear based on the write-up: the authors did not report undertaking any abstract screening, discuss the inclusion or exclusion criteria used or, in fact, seem to exclude any articles at all that appeared within its search, even if they were potentially irrelevant. Furthermore, the search and data analysis lacked depth, being restricted to a single database, a five-year search window, and only analysing metadata (e.g. authorship, citations) included with the Scopus search results. Sumiharsono et al.'s conclusions, therefore, give us only a very limited insight into the landscape of Character Education literature.

Secondly, Yolandini et al. (2023) carried out a bibliometric analysis focussing on Character Education research that had taken place in Indonesian elementary schools. Like Sumiharsono et al. (2023), they also analysed the most published institution, the most frequently published in journal, and the number of article citations. They found that: the Universitas Pendidikan Indonesia was the most published institution; Suriansyah was the most published author; the Journal of Physics Conference series was the most published in

journal, and the article with the most citations was 'Leadership Strategies of Principals, Teachers, Parents and Communities in Shaping Student Character' (Suriansyah, 2015). However, like Sumiharsono et al. (2023), the rigour involved in the review was unclear. The search term, database, abstract screening criteria, and inclusion and exclusion criteria were not identified in the article and, again, it is unclear if any articles were excluded from the search results. In addition, although the scope of the search was broader than that used by Sumiharsono et al. (2023), covering a 10-year period, this review is more narrowly focussed, only examining articles from a single country, and a single phase of education. This study's findings, therefore, also give relatively limited insight into the overall Character Education literature.

Thirdly, Muhtar et al., (2021) undertook a bibliometric analysis exploring the use of Character Education within Physical Education, searching the Scopus database for a nine-year period (2011-2020). From their retrieved articles they assessed the rate of publications, number of citations, article authorships, and frequency of collaborations. The number of publications per year related to Character Education and Physical Education was found to be steadily increasing, the number of citations per year to vary greatly over time, and around 10% of published institutions were found to have strong collaborative links with others. Unlike the other two bibliometric reviews, Muhtar et al. (2023) did apply inclusion and exclusion criteria to filter their original pool of articles, excluding a number of publications. However, like Sumiharsono et al.'s bibliometric analysis, the scope of the review was extremely narrow, focussing on only a single area of the curriculum, with a total of only 26 articles analysed. This means any conclusions drawn by the authors give only a narrow insight into the field of Character Education.

In summary, the most recently conducted bibliometric analyses of Character Education research lack rigour and/or focus on an extremely narrow slice of the field,

meaning they can only plausibly be considered to have scratched the surface of the complex, wide-reaching Character Education corpus.

Three Character Education meta-analyses have also been published in the past five years. The first of these is Brown et al.'s (2022) meta-analysis of Character Education programmes. This study sought to identify 'What Works' in Character Education, searching several educational academic databases for quantitative studies within K-12 settings that involved a control condition. The effect sizes for the Character Education interventions were found, on average, to be positive, although very few of the interventions found had been the subject of more than one study (Brown et al., 2022). This study, however, used an extremely broad definition of Character Education, describing it as '…any program intended to foster positive student development' (Brown et al., 2022, p. 121). This definition was used to inform the search terms used in the study, meaning that many of the programmes ultimately included in the review (e.g. Drug Abuse Prevention) would fall outside our dualistic definition of Character Education as developing positive traits and moral characteristics. In addition, the study only examined studies published up until 2017, so they do not take into account the most up-to-date research in the field.

The second Character Education meta-analysis was executed by Khadijah et al. (2021) and sought to evaluate the impact of schools' Character Education values on Mathematics outcomes in Indonesian schools. This review reported that, overall, Character Education had a positive effect on Mathematics achievement (Khadijah et al., 2021), but its findings are also of limited applicability to the field of Character Education as a whole, as they focused on just one academic subject in one particular country. A further methodological limitation of this review was that the authors only searched a single database (Google Scholar) to find relevant articles, giving them only a limited pool of data to analyse.

Jeynes' (2019) meta-analysis on Character Education's impact on student achievement and behavioural outcomes offers a large-scale review, drawing on 52 studies with over 225,000 participants, but suffers from methodological and interpretive limitations. The study quality is mixed, with a mean quality score of 1.88 out of 3, and many studies relied on passive rather than active control groups, failing to account for placebo effects and thus limiting causal interpretations. The inclusion of studies up to 2014, with an average publication year of 2004, also raises concerns about the currency of its findings, as it does not fully account for more recent developments in the literature. Additionally, the broad and somewhat vague definition of Character Education encompasses diverse interventions, obscuring which components are most effective. The analysis assumes homogeneity across programmes, despite differences in design, duration, and context, and the stronger effect sizes for behavioural outcomes over academic ones suggest potential confounding factors. The exclusion of qualitative studies further limits insight into the overall Character Education corpus, where it is possible that quantitative studies account for a minority of the literature. While the meta-analysis provides a useful overview of quantitative studies that existed a decade ago, it allows for very limited inferences, particularly given the exclusion of qualitative studies and the growing body of research over the past decade that has not been captured.

In summary, these meta-analyses provide some potentially useful indications of the positive impact that Character Education programmes can have. Nevertheless, Jeynes' (2019) and Khadijah et al's (2021) analysis both had narrow foci, incorporating only a small number of studies and specific outcomes, and although Brown et al.'s study (2022) was broader, it stretched its search for articles to beyond the field of Character Education research. In addition, by their very nature as meta-analyses, these reviews only analysed quantitative

studies which had reported some kind of effect size, therefore providing no insight into the substantial qualitative and non-empirical Character Education literature.

Concerns about Character Education Research

Since the emergence of Character Education as an important player in the field of education, a number of concerns have been raised as to the quality of the research conducted in its name.

Was et al. (2006) analysed the Character Education literature and identified a number of systematic weaknesses in the research methodologies used. The authors recognised that Character Education had already become embedded as a key element within educational policy, but that robust evidence was lacking as to the effectiveness of Character Education interventions, in particular, the lack of consistent use of robust, quantitative measures to assess desired Character Education outcomes. Was et al. (2006) ultimately concluded that it was 'essential' for researchers to robustly evaluate the effectiveness of particular Character Education programmes by conducting controlled, pre/post test designs interventions using effective measures.

In 2008, Slavin (2008) analysed the contemporary issues faced by researchers attempting to synthesise the educational programme evaluations to draw conclusion about effective practice, including for Character Education programmes. Their findings reinforced Was et al.'s (2006) conclusions, highlighting the importance of conducting randomised controlled trials to rigorously assess the effectiveness of Character Education interventions, and ensuring the publication of key statistics, such as effect sizes, to allow evidence to be able to be effectively aggregated, interpreted, and analysed. Slavin (2008) concluded by pointing out the importance of educators and researchers having the robust information they need to be able to apply Character Education in the real world.

This early, constructive criticism of Character Education research approaches from Was et al. (2006) and Slavin (2008) laid down the gauntlet to Character Education researchers to produce robust evidence to support the effectiveness of Character Education interventions. However, almost 10 years later, some of the same methodological criticisms were still being thrown at Character Education research. For example, Duckworth and Yeager (2015) raised concerns that measures being used in the assessment of non-academic programmes, including Character Education, were often poorly designed, lacking conceptual clarity, and called for the use of more robust, performance task-based measures in order to more effectively evaluate constructs such as virtues.

To date, 20 years on from these initial challenges/concerns being raised around the use of validated measures, as well as the importance being highlighted of effective pre/post test designs and reporting of key statistics, these have yet to be fully revisited across the Character Education corpus. Although the meta-analyses discussed in the preceding section all undertook some analysis of the quality of the papers included, this analysis was a) limited to the particular papers within their chosen samples and b) did not examine the measures employed by each study (Jeynes, 2019; Johnson et al., 2022; Khadijah et al., 2021). This continued lack of clarity as to the quantitative rigour of the research falling under the Character Education 'brand' is concerning for educators and researchers, both of whom need to be confident in the research upon which they are drawing to make decisions and implement Character Education programmes.

Purpose of The Current Review

Character Education has become a major player in the field of education. However, despite its impact, and the diverse and novel nature of its research corpus, currently there is no review available which gives an overarching view of the current publishing landscape of

Character Education research. Furthermore, despite a number of specific, methodological concerns having been raised about Character Education research over the past two decades, it is yet to de determined whether these have been addressed. The purpose of this scoping review is to tackle these two current gaps in the literature. This will be achieved by conducting a rigorous, overarching, time-unrestricted, multi-database review of the current publishing landscape of the peer-reviewed research falling under the Character Education brand. A further analysis will then be conducted on quantitative, longitudinal Character Educational studies to assess their methodological robustness in light of the previous criticisms.

Research Questions

RQ1: What is the current publishing landscape of Character Education research?

RQ2: How methodologically sound is Character Education quantitative research?

To address these questions, we initially searched the Scopus and Web of Science databases to generate a large, inclusive, time-unrestricted pool of relevant articles that use the term "Character Education". We then collated and analysed meta- (e.g. authorship), inferred (e.g. country of origin), extracted (e.g. use of control groups), and appended (e.g. Web of Science Journal Impact Factor) data for each article to answer each question in turn. The decision to use "Character Education" as the sole search term was not just a methodological choice but a deliberate effort to focus this review meaningfully. While fields like Positive Education and Social-Emotional Learning (SEL) share overlapping interests with Character Education, they are distinct in their aims and applications. Character Education stands apart as a cohesive field with its own conferences, audiences, research centres, and policy frameworks, operating largely independently of these adjacent domains. This also helps to address the critique offered above that previous reviews have been too encompassing, leading

to unmanageable heterogeneity and diluted findings. By focusing on articles that explicitly label themselves as "Character Education," this review strikes a balance: it ensures the inclusion of relevant studies from adjacent fields that see themselves as part of the Character Education corpus, while avoiding an overly broad search that would result in a dataset too heterogeneous to yield meaningful insights. This approach allows us to map the distinct terrain of Character Education with the precision it deserves, without overreaching too far into neighbouring, or indeed completely unrelated, territories.

Answering these research questions will, firstly, give clarity to both researchers and practitioners as to the current status of the research falling under the Character Education marque, identifying potential strengths and areas where more work is needed. Secondly, it will provide researchers and practitioners with critical information as to the methodological quality of quantitative studies carried out in the name of Character Education, particularly longitudinal ones, allowing them to make properly informed decisions as to how to enact Character Education in the real world.
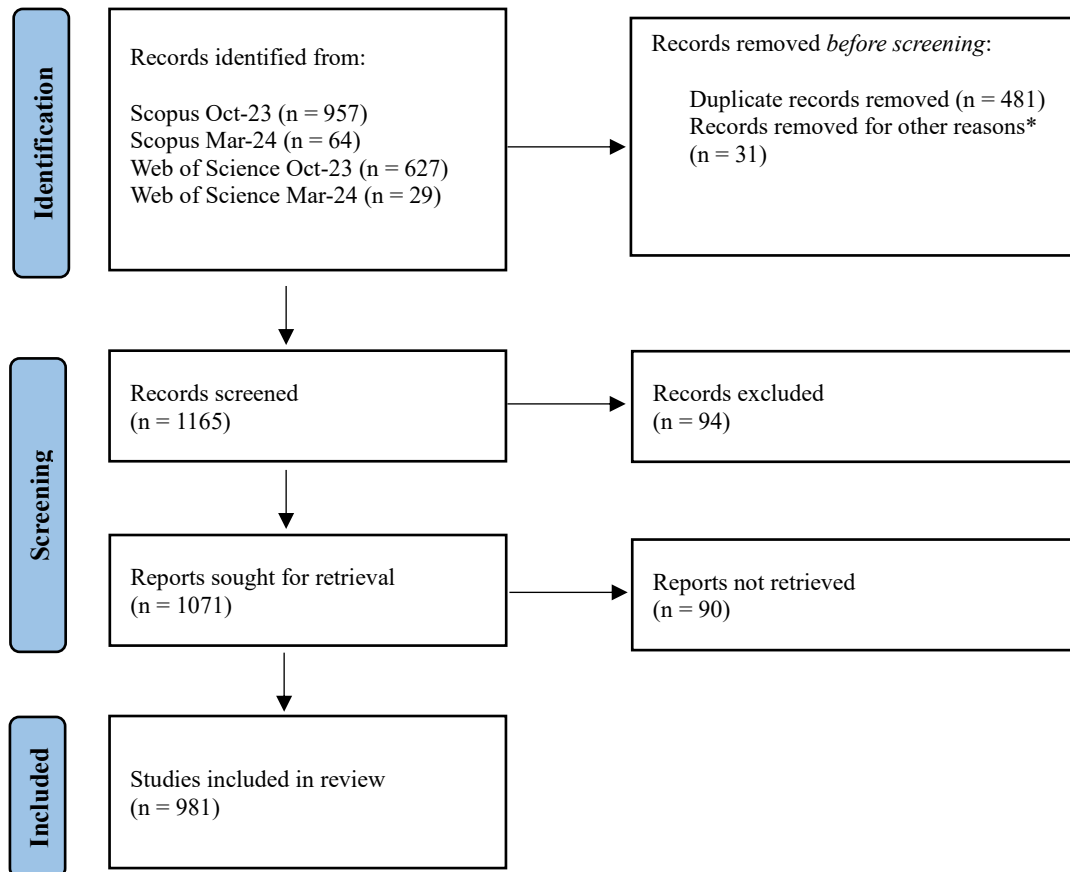
Methodology

In this scoping review, we followed the PRISMA guidelines for Systematic Reviews and Meta Analyses (Page et al., 2021). We initially searched the Scopus and Web of Science (WoS) databases for relevant publications in October 2023, and repeated the search in May 2024 to ensure the results were as up to date as possible. We used the single search term "Character Education" as our aim was to solely gather research falling under the Character Education brand. The inclusion criteria for the search were publications that were peer-reviewed, written in English, and no time limit was placed on the search.

Following our search, we screened the returned abstracts. We included articles that met the broad definition of Character Education, focussing on the development of positive and/or moral personal characteristics. Articles were excluded if they were irrelevant, for example discussing the teaching of Chinese character calligraphy. The abstracts were screened by the first author, with a 10% sample then checked by the second author. No discrepancies were found. Records were then sought for all remaining articles, with the final total of $N = 981$ records successfully retrieved and included in the study (Figure 1).

We then proceeded to aggregate the required data to answer our research questions. A number of required variables were already available as metadata from the initial database searches: *authorship, institution, funding, journal,* and *number of citations*. From the *institution* variable, we then inferred the *country/ies of origin* for each article. In order to provide insights into the robustness of the peer-review process to which each article had been subjected, we then retrieved the Web of Science Journal Impact Factors (*JIF*) for each journal and appended said JIF to the data set. The Web of Science JIF was chosen over the Scopus CiteScore as a measure of research quality as the Web of Science JIF is more up to date, using data from the previous two (Clarivate, 2024), as opposed to three (Elsevier, 2023) years.

Figure 1

PRISMA flow chart of included studies

Identification	Records identified from: Scopus Oct-23 (n = 957) Scopus Mar-24 (n = 64) Web of Science Oct-23 (n = 627) Web of Science Mar-24 (n = 29)	Records removed *before screening*: Duplicate records removed (n = 481) Records removed for other reasons* (n = 31)
Screening	Records screened (n = 1165)	Records excluded (n = 94)
	Reports sought for retrieval (n = 1071)	Reports not retrieved (n = 90)
Included	Studies included in review (n = 981)	

*book chapters, not articles or not available in English

A number of further variables were then extracted from the articles themselves. Firstly, the type of article. Articles were first coded as either *empirical* (gathered primary data) or *non-empirical* (did not gather primary data). The empirical articles were then further subdivided into *quantitative* (gathered only numerical data), *qualitative* (gathered non-numerical data) or *mixed* (gathered both numerical and non-numerical data). The *mixed* and *quantitative* articles were then analysed to identify a subgroup of *intervention* articles (studies which contained a quantitative, longitudinal component with, at a minimum, a post-test numerical measurement).

From the *intervention* articles, further data was then extracted for the following variables: *control group* (did they use a comparative group for the intervention?); *type of control group (*inactive – no treatment; active – placebo treatment; both – active and inactive); *randomisation* (did they randomly allocate participants to their groups?); *type of randomisation* (cluster – random assignment of groups of people; full – random assignment of individuals); reporting of *effect size* (Cohen's *d*, Pearson's *r* etc.), and reporting of *standard error*.

In order to analyse the robustness of the measures used in the *intervention* studies, the following further variables were extracted from these articles: *scale name* and *internal validation* (whether the study had conducted a factor analysis on the scale used as part of the study). The final variable, e*xternal validation,* of the scales was then conducted via a search of the literature to find if the scale had been previously subjected to factor analysis.

Initial coding for all variables was undertaken by the first author. The second author then randomly sampled 10% of the data for validation. A small number of errors were found in extraction of the *effect size* and *scale validation* variables. The first author, following additional training, then went back and checked these variables throughout the data set and made the required adjustments.

Once all the data had been extracted into Google Sheets, it was then transferred into an Amazon RDS Relational Database and then analysed using DBeaver v.24.02.24 (DBeaver, 2024) software with the help of AI SQL (SQLAI.ai, 2024).

Results and Discussion

For the sake of clarity, we have presented the results and discussion alongside one another, separated by research question.

RQ 1: What is the current publishing landscape of Character Education research?

Filed Overview

The number of articles published under the Character Education brand has increased exponentially over the past 30 years (Figure 2). In the 70 years between 1923 and 1993, a total of 121 articles were published under the Character Education marque. However, by 2013, the total number of Character Education articles had nearly tripled to 341. Then, from 2013 to 2023, an additional 618 further articles were added to the corpus – a remarkable increase in publication rate.

Figure 2

Cumulative frequency of Character Education articles
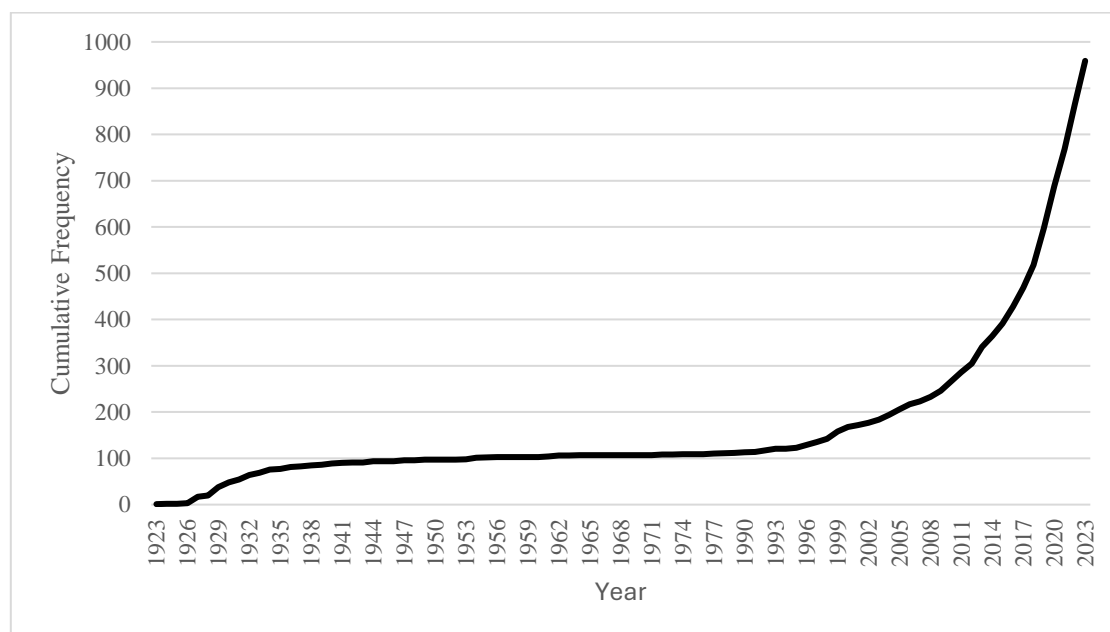
Such a rapid increase in Character Educations publications is an exciting development for those working in the field. Firstly, because it indicates the growing interest and importance of the field. Secondly, a greater volume of research being conducted gives the potential for more rapid discoveries and advances in the field. Thirdly, the more researchers engage in the field, the greater the opportunities for collaboration and potentially building links with other fields, and, lastly, a larger corpus raises the profile of Character Education, potentially opening up new opportunities for funding and sponsorship for future research.

In terms of what (or who) has been driving this significant increase in Character Education research output, an initial, significant rate increase can be seen in the early 1990s, followed by a much greater rise in publication rate at the starts of the 2010s. The initial rise in the 1990s can likely be attributed to the so-called "Return of Character Education" (Lickona, 1993) in the US. In 1993, the Character Education Partnership was launched, a national coalition determined to put Character Education at the top of the educational agenda (Lickona, 1993). This was followed in 1994 by the US Congress passing the "Improving America's Schools Act" which contained provision for grants to be distributed to education authorities to create and implement Character Education programmes. This generated many new opportunities for researchers to analyse the effects of Character Education in the real world, as well as creating a demand for robust theory to support these programmes, all driving a significant increase in Character Education publications.

The second publication rate increase in the early 2010s coincides with two major developments in the world of Character Education. The first was the founding of the Jubilee Centre for Character and Virtues in the UK in 2012, one of the first specialised research centres for Character Education in the world. In the last 12 years, the Jubilee Centre has become a major force in Character Education research, producing more publications than any other institution. The second major development was the introduction of Indonesia's national

"2013 Curriculum". This document stipulates that character must be taught throughout all stages of Indonesian education in line with local cultural values, with the aim of enhancing religious tolerance, empathy and mental health (Gunawan, 2017). The blanket implementation of Character Education across the fourth most populus country on earth (Source: *https://worldpopulationreview.com/countries,* accessed 20/11/24) created a plethora of opportunities for researchers to assess the impact of Character Education programmes in the real world, further driving up the rate of Character Education publications.

Geographical Distribution

The importance of these three countries in driving the exponential increase seen in Character Education publications in the last 20 years is reflected in the geographical distribution of Character Education publications (Figure 3). The US has contributed the most articles to the Character Education corpus (n = 365), followed by Indonesia (n = 297) and then the UK (n = 99). These three countries are, hence, the current powerhouse in the world of Character Education research, producing 77% of all research output.

However, the majority of research occurring in just three countries has two major potential implications. Firstly, the impacts of interventions can be culturally sensitive (Castro et al., 2010), meaning that if Character Education interventions have only been tested in a limited number of contexts, we may still be ignorant of their potential impacts in other settings. Secondly, the particular approaches to Character Education taken by these three nations, for example the more local values-based approach taken in Indonesia (Gunawan, 2017), will likely be disproportionately represented within the literature. This may lead researchers and practitioners, based on research volume, to believe that the Character Education approaches offered by these three major publications centres are the only valid

ones, and therefore may miss, or ignore, other just as valid approaches to Character

Education taken in other contexts.

Figure 3

Geographical distribution of Character Education research
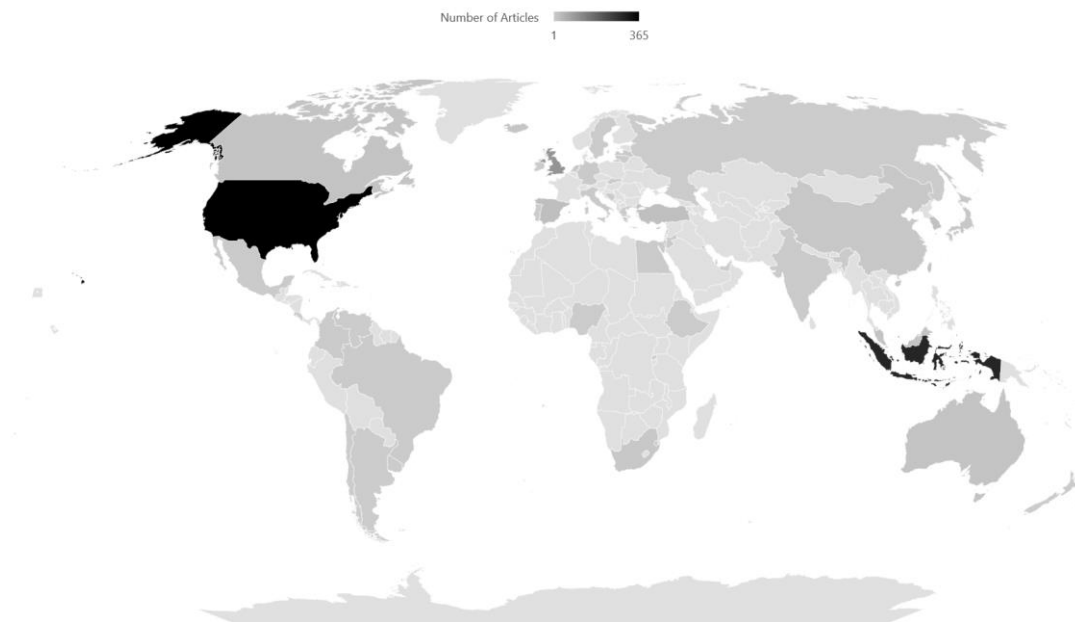
Nevertheless, although the US, UK and Indonesia account for 77% of all publications

carrying the Character Education marque, what is encouraging is the diversity of origin of the

remaining 23% of publications. In total, 54 different countries have contributed to the corpus,

with 1929 different authors from 803 different institutions publishing in more than 390

journals. This global nature of Character Education research is exciting for both researchers

and practitioners, giving the opportunity for Character Education to be theoretically and

empirically scrutinised in a multitude of contexts. Nevertheless, there are a number of caveats

to this excitement. Firstly, the modal number of publications per country is just one.

Secondly, around one fifth of countries did not receive any citations for their publications,

and thirdly, there is a risk that as Character Education has steadily shifted up the education

agenda, it just become the latest bandwagon that researchers feel they need to jump on to get

published, and so some of the work being produced may not be robustly grounded in the Character Education theory. Thus, whilst the global nature of Character Education research is encouraging, more in-depth, high-quality research is required in most contexts to fully understand the impacts of Character Education.

Types of Articles

Research published under the Character Education brand is split relatively evenly between theoretical and empirical articles (Figure 4). At first glance this is encouraging as it is important that theory can be used to effectively inform practice, and that empirical work can be used to test and hone theories (Bebeau et al., 1999). Nevertheless, in order for a Character Education particular theory to be robustly examined from multiple angles and in multiple contexts, more than one empirical study is likely to be required. From this perspective, it would appear that currently, Character Education research is rather heavy on the theory, but light on the testing.

Figure 4

Character Education article types
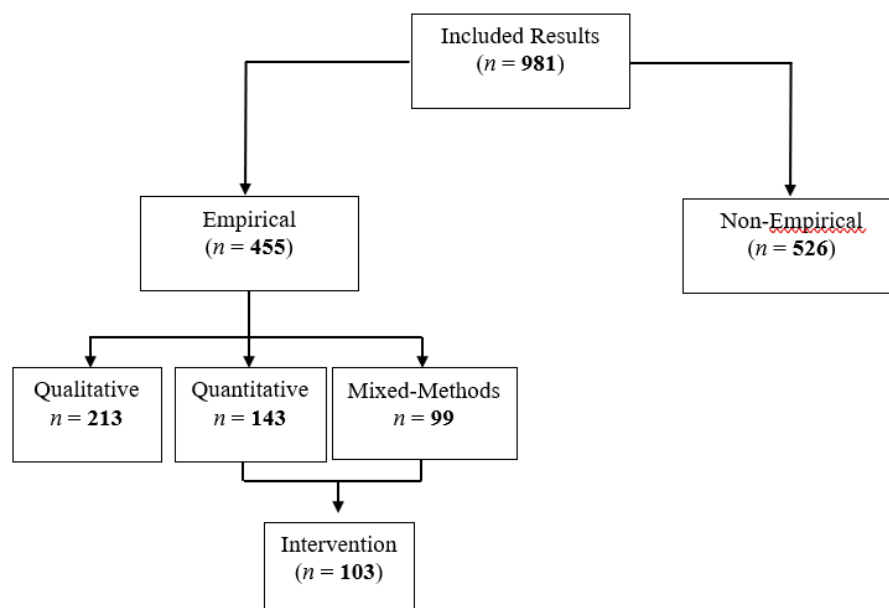
Within empirical Character Education research, qualitative research predominates, accounting for almost half of the research. However, although qualitative studies were the most common, they were also the least cited, receiving an average of only 4.0 citations per publication, versus 9.4 citations for quantitative articles. There are a number of possible explanations for this lack of cut through of qualitative research: a perceived lack of methodological quality in the current qualitative corpus; an actual lack of quality in the current qualitative corpus; or a perception within the field that qualitative research is somehow less valuable than other types of research. The implications of this may be that key insights are currently being missed/ignored from qualitative research into Character Education research or that more time/energy/money needs to be invested into qualitative research into the field to secure robust qualitative findings.

Intervention studies, arguably the key research type of empirical research for an ultimately applied field of academic study, accounted for less than a quarter of the empirical research. This is possibly a reflection of the greater expense and more time-consuming nature of such studies. Nevertheless, this relatively small number of intervention studies opens up the field to potential criticisms of a lack of real-world demonstration of its impact, potentially undermining its credibility as an effective educational intervention in the eyes of some.

Funding

153 different organisations have provided funding for research falling under the Character Education brand. The availability of so many different potential funding streams is great news for researchers in the field, as well as for practitioners, as greater investment in the field can only help but push forward our knowledge of Character Education. However, 116 of these organisations provided funding for only one article, indicating that although Character

Education may have broad appeal, not many organisations are willing to consistently fund research into the field.

By far the most frequent funder of research was the John Templeton Foundation (JTF), and its various subsidiaries, providing funding for around 5% of all Character Education research. The consistent funding provided by the JTF for Character Education research is a great boon to the field, but it does raise the question as to what extent the JTF's priorities shape the current Character Education corpus. This concern is reinforced by the fact that, on average, articles funded by the JTF received 46% more citations than those that did not. This higher number of citations may be being driven by the JTF having high standards for awarding its research grants, meaning the research it funds is of higher quality so has more cut through with researchers, or could arise from researchers receiving JTF funding feeling compelled to reference other articles that have previously been JTF funded. Either way, it is important for researchers and practitioners to bear in mind the influence the JTF's research priorities have had on the field when exploring Character Education research.

Research Quality

As with most things in life, quantity does not necessarily mean quality, and this is also the case in regard to research output (Haslam & Laham, 2010). The rapid increase in publications falling under the Character Education brand indicates the increased interest in conducting research into the field. However, there is a risk that this may, in part, be driven by researchers', previously mentioned, possible desire to leap on to the Character Education bandwagon to secure funding, or to appear to be up to date with current education trends. This big increase in research output, ergo, may have come at the expense of research quality. Although by answering RQ2: How methodologically sound is Character Education quantitative research? we will examine the methodological rigour of quantitative Character

Education research, it is worth at this point briefly examining the quality of the Character Education research output as a whole.

As previously elucidated, the current Character Education research comprises 981 articles written by 1929 authors from 830 institutions, publishing in 395 different academic journals. However, if we filter the results, removing data from any journals that did not have a 2022 Web of Science JIF (awarded to journals which, among other things, have a sufficiently robust peer review process in place (Clarivate, 2024)), these numbers are much lower: 609 articles, 1044 authors, 507 institutions, and 247 academic journals. More than a third of research appearing under the Character Education brand has, thus, not been subjected to the highest possible level of scrutiny, with almost half of Character Education authors only publishing in a journal without a Web of Science JIF.

All this has a particularly striking effect when analysing the most published individuals and institutions. Six of the top 20 institutions (Table 1), disappear from the list when the data is filtered to remove any publications from journals that do not have a 2022 Web of Science JIF. Interestingly, five of these hail from Indonesia; in fact, 222 of the Character Education articles which were published in a journal without a Web of Science JIF came from Indonesia. We posit three possible explanations for this. Firstly, as Indonesia is part of the Global South, it likely has access to lower levels of resources and academic expertise than more affluent nations, making it much more challenging to conduct robust research. Secondly, only 4% of Character Education articles from Indonesia were theoretical, meaning that the empirical articles being produced (which, ergo, is the vast majority) likely lack a strong, theoretical grounding. Finally, we noted that the level of English used in many of the articles emanating from Indonesia was poor, probably driven by the fact that English is not Indonesia's first language.

Table 1

Top 20 Most Published Character Education Institutions

	All Articles		Articles in WOSJIF+* Journals Only	
Rank	**Institution**	**Publication Frequency**	**Institution****	**Publication Frequency**
1	University of Birmingham	36	University of Birmingham	35
2	Universitas Negeri Yogyakarta	31	University of Missouri St. Louis	14
3	Universitas Pendidikan Indonesia	28	Universitas Negeri Yogyakarta	10
4	Universitas Negeri Malang	26	University of Edinburgh	10
5	Universitas Sebelas Maret	21	Universitas Negeri Malang	9
6	Universitas Negeri Jakarta	16	University of Alabama	9
7	University of Missouri St. Louis	15	University of Leeds	8
8	Universitas Negeri Makassar	12	Universitas Pendidikan Indonesia	7
9	University of Edinburgh	10	Wake Forest University	7
10	Boston University	10	Tufts University	7
11	University of Alabama	10	Boston University	7
12	University of Leeds	9	University of Pennsylvania	6
13	Universitas Negeri Semarang	9	University of Notre Dame	6
14	Universitas Ahmad Dahlan	9	Nanyang Technological University	6
15	Universitas Muhammadiyah Surakarta	9	University of Minnesota	6
16	Universiti Pendidikan Sultan Idris	9	Yale University	6
17	Wake Forest University	8	Universitas Sebelas Maret	6
18	California State University	7	University of Navarra	6
19	Tufts University	7	Universidad Internacional de La Rioja	6
20	University of Iceland	7	University of Iceland	6

*WOSJIF+ = Journal with 2022 Web of Science Journal Impact Factor

** Institutions no longer on this list: *Universitas Negeri Jakarta, Universitas Negeri Semarang, Universitas Ahmad Dahlan, Universitas Muhammadiyah Surakarta, Universitas Pendidikan Sultan Idris, California State University*

However, despite Indonesian Character Education struggling in some areas, even when filtered by Web of Science JIF, their institutions still make up four of the top 20 institutions for publications, and the empirical efforts made by Indonesian researchers to assess the impact and implementation of Character Education in the real world is

commendable, standing in contrast to some countries who have focussed much more heavily on theoretical study (e.g. in the UK, 63% of articles published were theoretical).

There are a number of implications to these findings. Firstly, it is important that when researchers examine the current Character Education corpus, they take into account the fact that a substantial chunk of research has not been submitted to the most robust possible review processes. Secondly, if Character Education's reputation is to be maintained, it is important to ensure that researchers focus on producing research of the highest quality. Finally, Indonesia clearly presents an amazing opportunity to conduct real-world research into Character Education, but to ensure this research is conducted in a robust fashion, local researchers would likely benefit from collaborations with institutions with expertise in both Character Education theory and study design.

In summary, Character Education research is a rapidly growing, pan-global phenomenon, driven forward, in particular, by the US, Indonesia and the UK, and backed by a multitude of funding organisations. However, the research is patchy in places in terms of quality, and this needs to be borne in mind by researchers and practitioners when exploring the field.

RQ2: How methodologically sound is Character Education quantitative research?

The methodological soundness of Character Education quantitative research will be evaluated by examining the use of scale validation, control groups, randomisation and statistical reporting in Character Education quantitative studies, in line with previous concerns raised about the field (Duckworth & Yeager, 2015; Slavin, 2008; Was et al., 2006).

Scale Validation

Quantitative studies published under the Character Education marque were deemed to have used valid measures if 50% or more of the measures used in the study had been

subjected to factor analysis, a prerequisite for the unambiguous mapping of theoretical constructs on to empirical phenomena (Watkins, 2018). Measure validation was assessed to be either internal (validated within the study itself), external (validated within at least one previous study) or both (validated both within the study and within at least one previous study).

The vast majority of Character Education quantitative studies did not use a validated measure (Figure 5), and only 3% of studies used a measure that had been subject to both internal and external validation. Possible explanations for the frequent use of unvalidated measures in Character Education studies may be a lack of availability of previously validated psychometric measures, or a lack of expertise amongst researchers in how to validate their own, novel measures given the interdisciplinary nature of the field.

Nevertheless, the lack of use of valid measures has grave implications. Without use of valid measurement tools, quantitative Character Education researchers cannot be confident in their findings. This calls into question many of the findings reported within the Character Education literature, and indicates that previous concerns raised about the use of unvalidated scales in the field have yet to be addressed (Duckworth & Yeager, 2015; Was et al., 2006). As Curren and Kotzee (2014) argue, character can be measured at a group level such that Character Education programmes can be evaluated, though measuring character at an individual level may be more challenging. If Character Education researchers do not make substantial and rapid moves in the direction of using validated measures, the field leaves itself open to serious criticism and possible loss of credibility.

Figure 5

Scale validation of quantitative Character Education studies
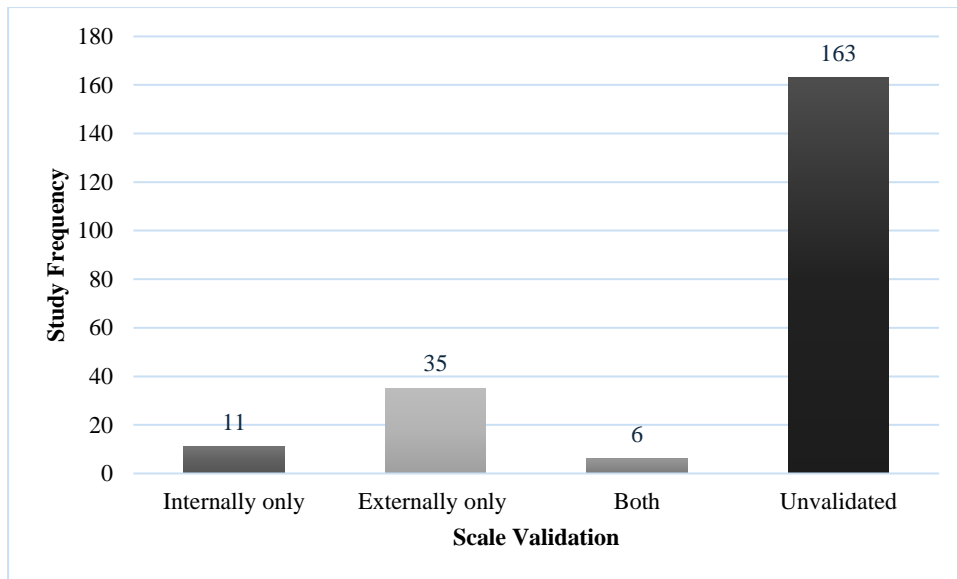
NB: Quantitative studies were not included if their sole purpose was scale validation or if they had not used a measure suitable for factor analysis.

Nevertheless, the lack of use of valid measures has grave implications. Without use of valid measurement tools, quantitative Character Education researchers cannot be confident in their findings. This calls into question many of the findings reported within the Character Education literature, and indicates that previous concerns raised about the use of unvalidated scales in the field have yet to be addressed (Duckworth & Yeager, 2015; Was et al., 2006). As Curren and Kotzee (2014) argue, character can be measured at a group level such that Character Education programmes can be evaluated, though measuring character at an individual level may be more challenging. If Character Education researchers do not make substantial and rapid moves in the direction of using validated measures, the field leaves itself open to serious criticism and possible loss of credibility.
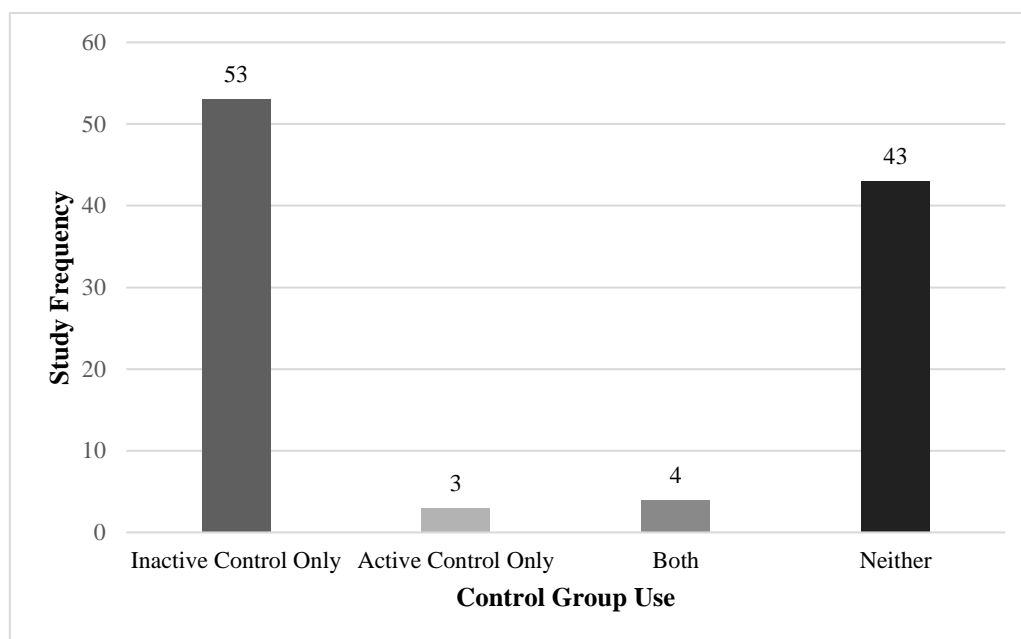
Control Groups and Randomisation

For intervention studies, inactive control groups allow researchers to test whether the intervention is more effective than simply the passage of time, and active controls allow researchers to test the intervention against an existing treatment whose efficacy has already

been established (Au et al., 2020). The purpose of randomly assigning participants to either type of control group is to ensure that the characteristics of the participants in each group are as similar as possible, avoiding selection biases (Stolberg et al., 2004).

In the Character Education literature analysed in this study, 58% of the intervention studies used some type of control group (Figure 6). However, only seven studies used employed an active control group. Of the studies which employed a control group, 70% of them did not employ any form of randomisation (Figure 7).

Figure 6

Frequency of control group use in Character Education intervention studies

The reasons for the low use of control groups and randomisation may be twofold. Firstly, practicality. Running interventions involving multiple groups is inevitably more complex, incurring a greater cost in both time and materials. It also clearly requires the availability of multiple groups of individuals, which researchers may not have access to. Secondly, expertise. Character Education is a field generally dominated by philosophers and educators,

who may not have the know-how to conduct and analyse research involving control groups and randomisation.

Figure 7

Frequency of control group randomisation in Character Education intervention studies
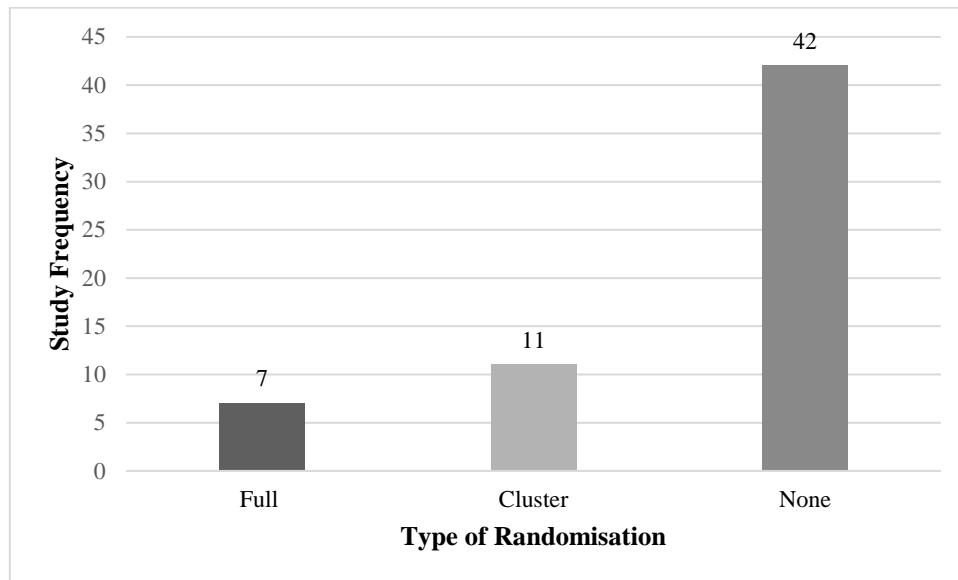
The significant numbers of Character Education intervention studies that did not employ a control group (particularly active controls), and the lack of randomisation used in the majority of studies that did employ a control group, are worrying for Character Education practitioners and researchers. For studies that did not use a control group, we cannot be sure if the intervention employed was any more effective than just the passage of time, or other alternative interventions. For the control studies that did not randomise, we cannot be confident that the results were not affected by selection bias. If we are to maintain confidence in Character Education as an educational intervention, these are issues that need addressing forthwith. They also, sadly, demonstrate the failure of many Character Education researchers

to heed Slavin's (2008) advice to ensure they gather the robust evidence which practitioners need to be able to effectively implement Character Education in the real world.
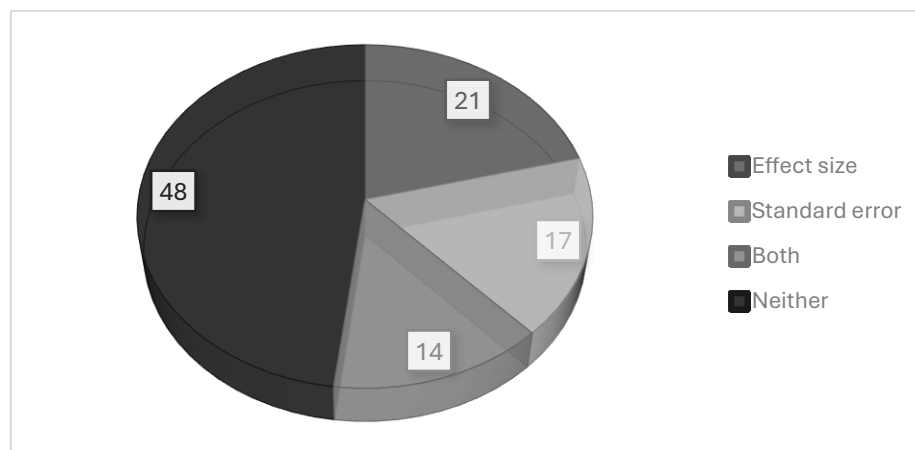
Statistical Reporting

Reporting of effect size and standard errors in intervention studies allows interpretation of the practical significance of any interventions tested (Eisenhart et al., 2006), and allows results of multiple such studies to be collated, for example using meta-analysis.

Almost half of intervention studies falling under the Character Education brand did not report an effect size or a standard error, and less than 15% of the studies reported both (Figure 8). As with the issues with measure validation, use of control groups and randomisation, this lack of reporting likely arises due to a lack of expertise in the field, in this case a lack of knowledge of how to interpret and evaluate the quantitative data produced from intervention studies.

Figure 8

Frequency of Character Education intervention study statistical reporting

The lack of effect size and standard error reporting seriously curtails researchers' and practitioners' abilities to understand the impact of interventions reporting in the large number of studies that lack one or more of these statistics. Furthermore, it leaves reviewers without

the key data that they require in order to conduct aggregative analyses on the field. All this, once again, indicates that methodological advice offered to the field of Character Education more than 15 years ago, in this case about the importance of reporting these statistics, has broadly gone unheeded (Slavin, 2008).

In summary, in answer to our RQ2: How methodologically sound is Character Education quantitative research?, the answer, despite the previous concerns raised and advice offered, (Duckworth & Yeager, 2015; Slavin, 2008; Was et al., 2006), seems to be – not very.

Future Directions and Limitations

In answering RQ1: *What is the current publishing landscape of Character Education research?,* we have established the rapidly increasing rate of publication of articles under the Character Education brand; the major role that Indonesia, the US and UK have had in driving up this rate; the pan-global nature of Character Education research; the lack of impact of Character Education qualitative studies; the relatively low number of intervention studies; the predominance of the JTF in funding Character Education research, and that a large chunk of Character Education research has been published in journals with less robust review processes. The implications of each of these findings have been discussed in the previous section, but each also opens up new opportunities and challenges for researchers publishing under the Character Education umbrella.

Firstly, more in-depth empirical study on Character Education needs to be conducted, particularly in contexts outside of the US, UK and Indonesia. It is critical that we understand the impact of Character Education interventions and, as these can be culturally sensitive (Castro et al., 2010), it is important that the effects of Character Education interventions are understood within all contexts in which it is being implemented. Secondly, more store needs to be placed in Character Education qualitative studies, a key factor in understanding the

deeper impact of educational interventions (Maxwell, 2012). This may entail researchers engaging more in the currently available Character Education qualitative research or helping to produce higher quality qualitative research output. Thirdly, researchers and practitioners need to be aware that a large proportion of the current Character Education corpus has not been subjected to the most rigorous review processes, and any examination of the current literature must be conducted with this in mind. Finally, countries that are currently producing less robust Character Education research, such as Indonesia, could benefit from collaborating with institutions in other countries with greater research expertise and access to resources.

In answering RQ2: *How methodologically sound is Character Education quantitative research?*, our findings show that: the majority of Character Education studies are being conducted without properly validated measures; control groups are being used in some cases, but active controls in particular are lacking; randomisation is occurring in only 30% of control group studies, and effect size and standard error are only rarely both being reported within Character Education intervention studies. These all reflect concerns that have previously been raised about Character Education research (Duckworth & Yeager, 2015; Slavin, 2008; Was et al., 2006), and require addressing rapidly if they are not to undermine the credibility of the field. We suggest the following actions need to be taken forthwith:

Firstly, future quantitative Character Education studies need to ensure they use validated Character measures. This will likely involve greater use of extant measures, as well as potentially the development of new ones to cover the conceptual blind spots of adjacent fields. This is a crucial first step to ensure that we have accurate insights into the impacts of Character Education. Therefore, while the field appears to be dominated by educators and philosophers producing theoretical or qualitative works, more resources should be allocated to psychometricians to strengthen this key area upon which any conclusions about the efficacy of Character Education depend.

Secondly, the methodological quality of quantitative intervention needs to be improved, with future studies ensuring they: (i) report key statistics, to ensure the impact of interventions can be understood and potentially subjected to meta-analysis; (ii) where possible, use randomisation to control conditions to minimise risks of selection bias and to control for potential confounding variables; and (iii) use *active* controls to help to rule out mere expectation/placebo effects within intervention studies.

The first limitation of this review was the search term used. Although we had good reason to use the single search term "Character Education", in doing so we will inevitably have missed some relevant research – for example, articles which used the key word "Character", but not "Character Education". To address this, a future study could be conducted using other closely-related search terms to Character Education to add further studies to our data set to provide an even deeper analysis of the field. The second limitation was that our search was restricted to English language articles only, ergo possibly missing articles from some research contexts. Our final limitation was that although this review presents insights into the "when, how, where" of empirical Character Education research, it does not address the "what" in terms of the topics of these articles. A future thematic analysis of the Character Education corpus could provide useful insights into the depth of research underpinning specific Character Education concepts.

Conclusion

Our study set out to answer the following novel research questions: RQ1: What is the current publishing landscape of Character Education research? and RQ2: How methodologically sound is Character Education quantitative research? We did so by conducting a rigorous, multi-database, time-unrestricted search for peer-reviewed literature falling under the Character Education brand, and then analysing meta-, inferred, appended and extracted data for each article in the data set.

In response to RQ1, we uncovered a big increase in the rate of Character Education publications in the last three decades, with thousands of authors from hundreds of institutions across the globe, supported by more than a 150 funders driving the field forward. However, we found this burgeoning volume of research to be heavily focussed in a small number of countries and, in places, to be beset by a lack of quality.

Addressing RQ2, we found many of the methodological concerns previously raised about the field, including concerning the use of valid measures, control groups, randomisation, and reporting of key statistics, to still be present. These methodological limitations need to be borne in mind by both researchers and practitioners when searching the current corpus, and these issues, along with the issues raised about general Character Education quality, need to be addressed practically by those in the field as soon as possible if Character Education is to sustain its current, exalted status within the field of education.

Declaration of interest statement: We have no conflicts to declare.

References

Arthur, J. (2005). The re-emergence of character education in British education policy. *British Journal of Educational Studies*, *53*(3), 239–254. https://doi.org/10.1111/j.1467-8527.2005.00293.x

Arthur, J., Moulin-Stożek, D., Metcalfe, J., & Moller, F. (2019). Religious Education teachers and character: Personal beliefs and professional approaches. https://www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/Research Reports/RE_Teachers_and_Character.pdf

Au, J., Gibson, B. C., Bunarjo, K., Buschkuehl, M., & Jaeggi, S. M. (2020). Quantifying the difference between active and passive control groups in cognitive interventions using two meta-analytical approaches. *Journal of Cognitive Enhancement, 4(*2), 192–210). Springer Nature. https://doi.org/10.1007/s41465-020-00164-6

Bebeau, M. J., Rest, J. R., & Navarez, D. (1999). Beyond the promise: A perspective on research in Moral Education. *Educational Researcher*, *28*(4), 18–26. https://doi.org/https://doi.org/10.3102/0013189X028004018

Brown, M., McGrath, R. E., Bier, M. C., Johnson, K., & Berkowitz, M. W. (2022). A comprehensive meta-analysis of character education programs. *Journal of Moral Education*, *May*. https://doi.org/10.1080/03057240.2022.2060196

Brunsdon, J. J. (2023). 'It's about bringing people together': One middle school Physical Education teacher's attempt at deploying character education. *Physical Education and Sport Pedagogy*, 1–13. https://doi.org/10.1080/17408989.2023.2256753

Castro, F. G., Barrera, M., & Holleran Steiker, L. K. (2010). Issues and challenges in the design of culturally adapted evidence-based interventions. *Annual Review of Clinical Psychology, 6,* 213–239. https://doi.org/10.1146/annurev-clinpsy-033109-132032

Clarivate. (2024). *Web of Science Journal Evaluation Process and Selection Criteria*. https://clarivate.com/academia-government/scientific-and-academic-research/research-discovery-and-referencing/web-of-science/web-of-science-core-collection/editorial-selection-process/journal-evaluation-process-selection-criteria/

Curren, R., & Kotzee, B. (2014). Can virtue be measured? *Theory and Research in Education*, *12*(3), 266–282. https://doi.org/10.1177/1477878514545205

Davies, I., Gorard, S., & Mcguinn, N. (2005). Citizenship education and character education: Similarities and contrasts. *British Journal of Educational Studies*, *53*(3), 341–358. https://doi.org/10.1111/j.1467-8527.2005.00299.x

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44*(4), 237–251. https://doi.org/10.3102/0013189X15584327

Eisenhart, M. A., Frederick, D., Grant, C. A., Green, J. L., Hedges, L. V, Felice, J., & Pellegrino, J. W. (2006). Standards for reporting on empirical social science research in AERA publications, American Educational Research Association. *Educational Researcher*, *35*(6), 33–40. https://doi.org/https://doi.org/10.3102/0013189X035006033

Elselvier. (2023). CiteScore TM metrics. https://www.elsevier.com/en-gb/products/copus/metrics/citescore

Gunawan, I. (2017). Indonesian Curriculum 2013: Instructional management, obstacles faced by teachers in implementation and the way forward. *128*(Icet), 56–63. https://doi.org/10.2991/icet-17.2017.9

Haslam, N., & Laham, S. M. (2010). Quality, quantity, and impact in academic publication. *European Journal of Social Psychology*, *40*(2), 216–220. https://doi.org/10.1002/ejsp.727

Improving America's Schools Act of 1994, 20 U.S.C.1513, 103rd Cong. (1994).

Jeynes, W. H. (2019). A meta-analysis on the relationship between character education and student achievement and behavioral outcomes. *Education and Urban Society*, *51*(1), 33–71. https://doi.org/10.1177/0013124517747681

Jordan, K. (2022). The feasibility of integrating insights from character education and sustainability education–a Delphi study. *British Journal of Educational Studies*, *70*(1), 39–63. https://doi.org/10.1080/00071005.2021.1897519

Khadijah, K., Suciati, I., Khaerani, K., Manaf, A., & Sutamrin, S. (2021). Schools' character education values and students' Mathematics learning achievement: A meta-analysis. *Cakrawala Pendidikan*, *40*(3), 670–683. https://doi.org/10.21831/cp.v40i3.39924

Kristjánsson, K. (2015). *Aristotelian Character Education*. Routledge.

Lickona, T. (1993). The return of character education. *Educational Leadership*. https://www.scirp.org/reference/referencespapers?referenceid=161647

Maxwell, J. (2012). The importance of qualitative research for causal explanation in education. *Qualitative Enquiry*, *18*, 665–661. https://doi.org/10.1177/1077800412452856

Ministry of Education, S. (2020). *Character & Citizenship Education (CEE) Syllabus*. https://www.moe.gov.sg/-/media/files/secondary/syllabuses/cce/2021-character-and-citizenship-education-syllabus-secondary.pdf

Muhtar, T., Supriyadi, T., Lengkana, A. S., & Cukarso, S. H. I. (2021). Character education in physical education learning model: A bibliometric study on 2011-2020 scopus database. *International Journal of Human Movement and Sports Sciences*, *9*(6), 1189–1203. https://doi.org/10.13189/saj.2021.090613

OCDSB. (2024). *Character Education*.
https://www.ocdsb.ca/our_schools/character_education

Ofsted. (2019). *The Education Inspection Framework*.
https://www.gov.uk/government/publications/education-inspection-framework/education-inspection-framework-for-september-2023

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *The BMJ*, *372*. https://doi.org/10.1136/bmj.n71

Slavin, R. E. (2008). Perspectives on evidence-based research in education—what works? Issues in synthesizing educational program evaluations. *Educational Researcher*, *37*(1), 5–14. https://doi.org/10.3102/0013189x08314117

Stolberg, H., Norman, G., & Trop, I. (2004). Randomized Control Trials. *American Journal of Roentgenology*, *183*. https://doi.org/10.2214/ajr.183.6.01831539

Sumiharsono, R., Safrudin, M., & Ramadhan, S. (2023). *Research Trends On Character Education Based On Scopus Database From 2018 To 2023: A Bibliometric Analysis*. https://doi.org/10.30868/ei.v12i02.4085

Suriansyah, A. (2015). Strategi Kepemimpinan Kepala Sekolah, Guru, Orang Tua, Dan Masyarakat Dalam Membentuk Karakter Siswa, *Jurnal Cakrawala Pendidikan, 2*(2), 234-237. https://doi.org/10.21831/cp.v2i2.4828

Was, C. A., Woltz, D. J., & Drew, C. (2006). Evaluating character education programs and missing the target: A critique of existing research. *Educational Research Review*, *1*(2), 148–156. https://doi.org/10.1016/j.edurev.2006.08.001

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/https://doi.org/10.1177/0095798418771807

Yolandini, B., Suabuana, C., Muhammad, I., & Triansyah, F. A. (2023). Analysis bibliometric: Character Education research in elementary schools on one decades. *JIIP - Jurnal Ilmiah Ilmu Pendidikan*, *6*(7), 5485–5492. https://doi.org/10.54371/jiip.v6i7.2582