

‘It made me think how I should treat others and how I should help people who need it’

The complexities of exploring the impact of Holocaust education

Rebecca Hale

UCL Centre for Holocaust Education, UCL Institute of Education, London

r.hale@ucl.ac.uk

Abstract

It is generally agreed that learning about the Holocaust has some impact on students. However, discussions about the nature and magnitude of impact tend to be intuition-based rather than evidence-based. This is exacerbated by studies giving insight into how Holocaust education is related to salient variables rather than studies which demonstrate that teaching and learning about the Holocaust directly results in particular outcomes. Additionally, studies have drawn on small samples or presented findings which rely only on participants’ own perceptions of impact rather than actual impact measurement. These issues have contributed to a muddled and contestable narrative about impact in this field. This paper explores the complexities of studying the impact of Holocaust education, considers the feasibility of using randomised control trials, and appeals to those working in the field to be mindful of making and accepting spurious claims about impact.

Key words

Holocaust education, impact, schools, teaching, learning

Introduction

The quote in the title of this paper was from a 14-year old student spoken after spending several weeks learning about the Holocaust in her history lessons. Sentiments like this, where

students reflect on what the Holocaust means to them and the attitudes and actions it galvanises, are probably not unusual. However, studies which explore the nature and magnitude of this impact in a systematic way are few and far between and leave many questions unanswered. For example, it remains unknown whether students' good intentions – in this case, to 'help people who need it', become concrete actions. It is also unclear how far Holocaust education instils new attitudes and behaviours or taps into existing inclinations. That is to say, arguably this student would have already known she should help people in need and thus, rather than producing new attitudes or behaviours, learning about the Holocaust affirmed what she already thought.

In England, the government annually provides millions of pounds of funding to Holocaust education organisations and initiatives. In 2015, they commissioned a new Holocaust Memorial and Learning Centre, due to be built next to the Houses of Parliament, with an initial contribution of £50 million from the government (GOV.UK, 2018) and a further £25 million committed in 2019 (GOV.UK, 2019). The scale of investment and interest in Holocaust education and memorialisation inevitably invites questions from those inside and outside government about purpose and impact. Additionally, the government's 'what works' agenda has put the issue of impact into sharp focus and has contributed to policy makers' increasing interest in a medical model of research and championing experimental methods like randomised controlled trials (RCTs) (Whitty & Wisby, 2016). This is most notable in education with the establishment of the Education Endowment Foundation (EEF) which has committed over £96 million to test the impact of 160 educational programmes (EEF, 2018). In view of these developments, increasing calls for evidence showing the impact of Holocaust education, particularly through the use of RCTs, are arguably inevitable.

RCTs have been described as the 'gold standard' in evaluation research (Concato, Shah, & Horwitz, 2000; D. J. Torgerson & Torgerson, 2008). However, despite endorsement of this

approach from academics, practitioners and policy makers, it has been argued that RCTs are less suitable for evaluating the impact of complex multi-layered interventions (Cupitt, 2015). The less standardised an education programme is and the less able the researcher is to tightly control salient variables, the trickier it becomes to conduct an RCT which will produce valid and reliable findings (Sullivan, 2011).

Teaching and learning about the Holocaust is a complicated, multifaceted and exacting process, which calls on teachers and students to grapple with disturbing, challenging and nuanced content. Thus, Holocaust education programmes certainly fit the description of being complex and multi-layered. This paper will focus on the challenges and considerations involved in evaluating the impact of Holocaust education programmes; issues requiring urgent and judicious attention because the subject is taught in 135 countries (Unesco, 2014), often with considerable funding and time invested. The paper will first present an overview of RCTs and compare them to quasi-experimental research and qualitative approaches. Particular attention is given to RCTs because of increasing endorsement of this approach by researchers and policy makers. Undoubtedly, the characteristics and issues related to RCTs and the other research methods identified are complex, and the scope of this paper precludes thorough examination. However, this overview will give some context when considering the possible ‘outcomes’ for teaching about the Holocaust. In particular, two domains of outcome will be discussed – attitudes and historical knowledge – with reference to widespread teaching aims used in Holocaust education, as well as existing research exploring impact. This paper will conclude with reflection on the feasibility of applying RCTs to Holocaust education, highlight the important role of qualitative research, and appeal to researchers and educators working in the field to be more cognisant and nuanced in how they interpret, discuss and present ‘evidence’ of the impact of Holocaust education.

Evaluating education programmes

In RCTs, participants are randomly allocated to either the experimental group where they are exposed to an intervention or the control group where they do not receive it. A comparison group may be used where participants receive an alternative treatment. Random allocation means overall the groups will have the same characteristics as each other; all known and unknown characteristics that might affect the results are present in all groups and will be cancelled out in the analysis (D.J. Torgerson & Torgerson, 2008). Thus, any observed group differences can be reasonably attributed to the intervention (Cook, 2002).

There has been considerable resistance to using RCTs in schools (Connolly, Biggart, Miller, O'Hare, & Thurston, 2017; Cook, 2002; Slavin, 2002). Reasons include, the challenge of recruiting schools willing to be randomly assigned to a group (even if that means being designated as a control school) and ethical concerns about depriving students of a programme thought to be beneficial (Slavin, 2002). Connolly et al., (2017) also highlighted concerns that RCTs symbolise the old positivist tradition, utilising artificial laboratory-style approaches to attempt to control educational settings in a similar way to how natural scientists control variables in experiments. This, they argue, is at odds with the complexity of beliefs, behaviours and interactions that play out in the school environment.

A related, though arguably much less robust approach, is the quasi-experimental design where outcomes from intervention and control groups are compared without using random allocation to assign the groups (Cook 2002; D. J. Torgerson & Torgerson, 2008). One example is the prepost with non-equivalent control group design (Handley, Lyles, McCulloch, & Cattamanchi, 2018). A school opts to implement an education programme (the intervention group) and their results are compared to a 'control' school that is not implementing the intervention (with researchers sometimes targeting schools perceived to

share salient variables with the intervention school). Quasi-experiments tend to introduce additional variables that influence the outcomes being measured. For example, the intervention group may be motivated in some way for the education programme to succeed (Cupitt, 2015). Despite the criticisms levelled at quasi-experimental studies, they are prevalent in educational research (Coe, 2009; Walser, 2014). They provide a means of exploring the impact of interventions where random allocation is deemed not feasible, and there are a number of quasi-experimental designs that can be utilised to increase validity (Handley et al., 2018).

Both RCTs and quasi-experimental methods have been critiqued for not giving insight into the causal processes that give rise to particular outcomes. Maxwell (2012) (pp. 658-659) argued:

The idea that randomised experiments or structural equation models can provide valid general conclusions about the effect of an intervention, in the absence of any understanding of the actual causal processes in which these processes were situated, or the meaning that the intervention and contexts had for the participants, is an illusion. We need qualitative methods and approaches in order to understand “what works” and why.

There are a number of strategies drawing on qualitative data which can provide methodologically valid means of exploring causation. This includes: the collection of rich data to generate, develop and test theories to account for the ‘outcomes’ of a particular programme; conducting narrative and connecting analysis to categorise, compare and identify connections between events and processes in a specific context; and the triangulation of data from a range of individuals or settings utilising different research methods (Maxwell, 2004).

In recent years the EEF have highlighted the importance of implementation and process evaluations (IPE). IPEs examine how an intervention is put into practice, how far it achieves its intended outcomes, and the factors that influence these processes. While RCTs look at ‘what works’, IPEs enable researchers to understand *how* and *why* interventions work (or how and why they did not work) (Humphrey et al., 2016). In this way, the feasibility of using RCTs to evaluate complex multi-layered interventions is strengthened. Oakley, Strange, Bonell, Allen and Stephenson (2006) provide such an example; they evaluated a peer-led sex education programme using an RCT which was augmented by exploring the implementation, receipt and context of the programme using questionnaires, focus groups, interviews, researcher observations and field notes. Thus, the systematic and meaningful utilisation of process evaluations integrated with RCTs to explore the impact of complex education interventions provides a promising avenue for research looking at the impact of Holocaust education programmes.

Aims of Holocaust education

When exploring the impact of Holocaust education, educators first need to reflect on what they want students to achieve after learning about the Holocaust and how they will determine if these aims have been met. However, Biesta (2015, p. 83), cautions:

The predicament here is whether we are measuring and assessing what we consider valuable, or whether bureaucratic accountability systems have created a situation in which we are valuing what is being measured, i.e. a situation where measurement has become an end in itself rather than a means to achieve good education in the fullest and broadest sense of the term.

In terms of Holocaust education, certainly in England, this taps into some particularly salient issues. Since 1991 it has been mandatory for students to learn about the Holocaust as part of

the history National Curriculum at Key Stage 3 (when they are aged 11-14 years)ⁱ. It was introduced into a crowded history curriculum without clear rationale or guidance; a situation which prevails today as the government has continued to avoid articulating what it wants teaching of the Holocaust to achieve (Pearce 2017).

Arguably, this situation gives teachers the opportunity to deliver Holocaust education programmes with content they consider valuable and not be constrained by prescriptive guidance or what can be aligned with (and measured for) particular agendas. However, the reality is much more complex and emotive. For instance, Pearce (2017) points out that at the turn of the millennium the Labour government was referring to the “lessons” of the Holocaust and oversaw the enactment of Holocaust Memorial Day and its directive to teach, learn and remember the Holocaust and its “lessons” to tackle prejudice and prevent other atrocities taking place. While such endeavours are well intentioned, as Bauer (2016, ix) argues ‘to distil the complexity of the Holocaust into a moral fable is bound to lead to oversimplification, half-understandings, and superficial knowledge’.

Research with teachers has provided evidence of a similar “lessons” discourse in schools. Amongst history teachers, the most common teaching aims of Holocaust education were to develop students’ understanding of prejudice, racism and stereotyping in society (67.7%) and encourage them to learn lessons from the Holocaust so that a similar human atrocity never happens again (53.6%). The levels of endorsement recorded for these non-historical aims were not achieved for the historical-based aims, with a quarter of teachers (24.9%) selecting ‘to deepen knowledge of World War II and Twentieth Century history’ and 18.7% selecting ‘to understand and explain the actions of people involved in and affected by an unprecedented historical event’ (Pettigrew et al., 2009). Therefore, despite being identified as a compulsory subject on the history National Curriculum, teaching aims related to the historiography of the Holocaust were often not prioritised.

In England, the Holocaust is also covered (on a non-statutory basis) in other subjects, including religious studies, citizenship and English (Foster et al., 2016; Pettigrew et al., 2009). The most frequently cited teaching aims for teachers of these subjects mirrored the most frequently cited aims for history teachers (Pettigrew et al., 2009). Thus, regardless of subject, teaching aims which speak to broader social agendas are prevalent. This has implications for how students conceive of and understand the Holocaust, which in turn has implications for how the impact of Holocaust education is conceptualised and measured (this will be discussed in the next section).

The utilisation of a range of teaching aims reflects a longstanding debate about what the purpose and focus of Holocaust education should be. This debate is complex, and it is not possible to provide a full account in this paper (for further discussion see: Feinberg & Totten, 2016; Kinloch, 1998; Lindquist, 2008; Pettigrew, 2010; Pettigrew, 2017; Salmons 2010). However, it is helpful to be familiar with two different perspectives often used as a shorthand approach to frame the debate: (a) the aim(s) of teaching about the Holocaust should relate to developing specific historical understandings which potentially offer perspectives on the contemporary world and (b) the aim(s) of teaching about the Holocaust should relate to social, moral and/or civic outcomes.

Kinloch (1998), a proponent for teaching about the Holocaust as *history*, argued teaching approaches which seek to present cautionary messages about discrimination and warn young people not to participate in or be indifferent to future genocides are facile because students will make their own moral and social judgements which are unlikely to be affected by their teachers. Karn (2012) illustrated this when recounting an experience with his students. He found attempting to imprint a superficial moral 'lesson' by connecting the past to the present unwittingly risked students rejecting this view rather than adopting it.

However, historical aims are not necessarily incongruent with the ‘non-historical’ aims teachers draw on. Educators and academics have argued for the importance of students developing secure historical knowledge about the Holocaust to then be supported to critique and reflect on what happened (Feinberg & Totten 2016; Foster et al., 2016; Karn, 2012).

Thus, while there are important alternative curricular and disciplinary framings of the Holocaust, secure historical knowledge is a fundamental basis from which to draw other forms of understanding (Foster et al., 2016).

Grouping teaching aims as simply ‘historical’ or ‘moral’ is also problematic because it does a disservice to the numerous and intricate rationales and aims which teachers draw on in their practice. Similarly, it is unhelpful when non-historical aims are referenced as ‘moral’ aims, becoming an erroneous catch-all term for ‘moral’, ‘social’, ‘civil’, ‘emotional’/’affective’, and/or ‘spiritual’ aims (Pettigrew, 2017). This all-inclusive approach masks important differences between these concepts, and for both educators and researchers, hinders interpretations of the many and varied ways that Holocaust education is likely to impact on young people.

Another problematic issue pertains to how the relationship between learning about the Holocaust and achieving socially orientated aims has at times been presented as occurring almost by osmosis. That is to say, exposure to the historical details of the Holocaust automatically leads to students becoming more moral and tolerant individuals and developing attitudes aligned with accepted narratives of right and wrong (Levy & Sheppard, 2018; van Driel, 2003). Thus, the mechanisms which enable this process are not only unknown, but also unquestioned.

Given the discourse surrounding teaching aims, it appears logical that impact research should explore both students’ historical knowledge about the Holocaust and non-historical variables

which tap into moral, social, civil, emotional/affective, and/or spiritual aims. The challenge for educators and researchers will be to identify and define salient variables, and then to isolate and measure them.

Impact research in Holocaust education

Only a very small number of studies have sought to explore the question of ‘impact’ in Holocaust education and even fewer have done this systematically. Instead, there has been a tendency in the field to make spurious claims about impact based on data which can only give insight into associations, conduct studies using small sample sizes, and/or present findings based on participants’ perceptions of impact (rather than actual impact).

The importance of differentiating between data which can show the *impact* of one variable on another variable and data which shows a *relationship* between variables cannot be overstated. For example, in a study by Jedwab (2010) the data showed relationships between variables, not causal links. Thus, the sentence ‘the importance of Holocaust knowledge as a causal factor in favourable views of diversity is borne out to some degree by the results of the survey’ is ambiguous (Jedwab, 2010, p. 284). The data were collected at one point in time and therefore it is unknown whether Holocaust knowledge caused favourable attitudes or whether favourable attitudes motivated individuals to learn more about the Holocaust.

Holocaust education remains an under-researched field. Therefore correlational and descriptive studies such as Jedwab (2010) provide useful starting points about potential salient variables that, if found to be related to Holocaust education, could then be examined more systematically in an RCT. However, it is essential for generalisations about causality to be avoided and data showing associations to be clearly presented as such. Even where RCTs are used, claims of impact should be appropriately nuanced and tentative (Connolly et al., 2017).

Another example demonstrating the caution required when measuring, describing and claiming impact is provided by Eadson, Pearson, Foden, Sanderson and Bashir (2015). They reported that taking part in Holocaust Memorial Day (HMD) had an impact on adults' knowledge and understanding, and led to changes in behaviours and attitudes (for example, participants reported being more likely to stand up to unfair treatment of others). However, while the sample size was large, the data were based on the respondents' reflections (for instance, how well informed about the Holocaust they felt before and after taking part in HMD) not actually concrete information about what they had learned, and no pre-test data and/or control group were used for comparison. The sentence (page 7) 'There is clear evidence that people's knowledge and understanding of genocides has been increased as a result of HMD' is especially misleading. While it is possible the HMD activities had an impact on the participants, the type of data collected did not provide 'clear evidence' for the claim being made, thus making the impact statement untenable.

A study by Gross (2017) evaluated an education programme where Holocaust survivors and students were brought together in interactive workshops. Details about the impact of the programme were ambiguous, for example (page 600) 'students didn't just learn history; they learned about the strength of the human spirit'. Thus, the conclusions lacked clarity about what the data could reliably show. Notably a survey sent to 117 students was only completed by 24 of them, giving a small and biased sample who were particularly motivated to give their feedback.

A small number of Holocaust education studies have utilised comparison and/or control groups. For example, Cowan and Maitles (2007) conducted a longitudinal study and found that primary school children were less likely to think it was acceptable to make racist comments about different groups after they had learned about the Holocaust, and this remained the case when they were followed up at secondary school. Moreover, when at

secondary school these students were less likely to think that racism had nothing to do with them compared to a group of their peers who had not learned about the Holocaust in primary school.

Gordon, Simon and Weinberg (2004) used two experimental groups of undergraduate students who participated in either a course exclusively about genocide and the Holocaust or a course about the Holocaust in the context of the Second World War. The control group were students learning about American politics with no discussion of the Holocaust.

Knowledge about the Holocaust improved from before to after participation for the students who learned exclusively about genocide and the Holocaust. In the other experimental group and the control group there were no significant differences in knowledge from before to after the course. Across all groups there were no significant differences from before to after the courses with regards to levels of antisemitism. Although, levels of antisemitism were already low to begin with so there was little scope for any sizeable positive changes in these attitudes.

The use of intervention and control groups in these two examples show the value of comparisons, however the groups were not randomly allocated. The studies also demonstrate how the magnitude of change observed in impact studies needs to be scrutinised for what it means in the 'real world'. For instance, Cowan and Maitles's (2007) found the proportion of children agreeing it was acceptable to make racist comments about Jews decreased from 4.6% to 3.1% after learning about the Holocaust and further decreased to 2.3% at follow-up. However, these changes were incremental and given the sample size, equate to no more than a few students changing their attitudes.

While attitudinal change even in just a small number of students is patently important, educators should reflect on how far a programme is addressing attitudes, skills or knowledge young people already possess, and instead look to challenge and develop students beyond

their pre-existing notions. It is also important to reflect upon the meaning of numerical data and statistical analyses in the broader 'real world' context. Results which are statistically significant are not necessarily significant when viewed in real terms. Indeed, reliance on statistical significance to interpret results is increasingly coming under scrutiny (see Gorard, 2010).

To date, a study by Barr et al., (2015) appears to be the only published RCT into a programme which teaches about the Holocaust. In this RCT, conducted in the United States, schools that had no or extremely limited exposure to the Facing History and Ourselves programme were randomly assigned to either the intervention condition (where teachers participated in Facing History professional development and then delivered the programme to their studentsⁱⁱ) or the "as is" control condition (where teachers continued to use their standard practises for history and language arts courses). Despite some participant attrition from baseline to follow-up, 113 teachers and 1,371 students took part in the follow-up surveys. The results showed the programme had a positive impact on ten of the twelve teacher outcomes measured including self-efficacy and satisfaction with professional support, engagement and growth. For students, five out of fifteen outcomes measured were found to be significantly impacted upon by the programme (overall historical understanding, political tolerance, civic efficacy, opportunities to engage in civic matters in class and open classroom climate teacher practices), although the effect sizes were relatively small.

Overall, the programme seemed to have less impact on young people, including their civic responsibility and civic participation. This could be because the instruments were unable to capture this sort of impact rather than the programme not having much influence on students. Even so, the programme appeared to have more impact on teachers. This was perhaps unsurprising because teachers directly received the programme in its "purest" form by participating in the Facing History summer seminar and having at least one planning meeting

with Facing History educators. Whereas, the students received the programme through the lens of their teachers' knowledge and interpretations of the programme which might have deviated from the 'pure' form. Indeed, Barr et al found the fidelity of the programme varied with almost half of teachers adhering to the programme's time and resource expectations with high fidelity, and the remainder with medium or low fidelity. Collectively, these issues demonstrate both the possibilities and complexities of using RCTs in Holocaust education, and highlight the importance of researchers reflecting on what can credibly be concluded about impact and what should be interpreted more tentatively.

The challenges of defining and measuring impact in Holocaust education

In view of teachers' proclivity for including 'non-historical' teaching aims, efforts to study and understand attitude change are well-reasoned. However, this approach presents a challenge because measures that reliably tap into attitudinal concepts like prejudice are difficult to create and often fail to capture the real impact of teaching, especially when this is expected to manifest in attitude changes or actions in the future (Pettigrew et al., 2009).

There is a large body of psychological theory and research which demonstrates that predicting and changing attitudes and behaviours is a complex process. For instance, Fishbein's and Ajzen's (2010) reasoned action model argues that attitudes, perceived norms and perceived behavioural control (an individual's perception of their capability to perform a behaviour) influence an individual's behavioural intention, and the stronger their intention the more likely it is the behaviour will be carried out. However, while perceived control might be high, a lack of actual control through lack of skills or the presence of environmental constraints can prevent people from acting on their intentions. In the context of Holocaust education, this model suggests that after learning about the Holocaust a student might be galvanised to challenge prejudice and discrimination in their community. They may believe

the consequences of doing this are positive; however the student may also predict that as an adolescent they will have little sway upon the views of those in their community. So, while Holocaust education had an initial impact on this student's behavioural beliefs, the longer term impact on their attitudes and/or intentions cannot yet manifest. This shows how complex and prolonged the process of attitudinal and behavioural change is, and that researchers could break down some of these dimensions and, for example, look at how Holocaust education influences different kinds of beliefs.

For those advocating historical teaching aims, the impact of Holocaust education on students' historical knowledge will be of interest. To measure this, researchers can create questions to assess what respondents know (for example, see Foster et al., 2016; G. K. Starratt, Fredotovic, Goodletty, & Starratt, 2017) or they can ask respondents to self-assess their level of knowledge (for example, see Cowan & Maitles, 2007; Jedwab, 2010). There are strengths and weaknesses to both approaches. For instance, using knowledge-based questions alludes to expectations about what students should know (Chapman & Hale, 2017). However, there is no consensus around a common set of questions to establish benchmarks to indicate satisfactory knowledge (Jedwab, 2010). While measuring respondents' self-assessment of their knowledge can address this issue, students' perceptions of their knowledge are not always accurate. This was demonstrated by Foster et al. (2016) in numerous instances where students were very confident the answers they gave to survey questions were correct, when in fact they were incorrect. This was particularly stark when asked what happened when a member of the military or police refused an instruction to kill Jewish people. Two thirds of students incorrectly said the military or police would be shot for refusing to obey an order, yet 67% of them were 'fairly confident' or 'very confident' this was the correct answer.

Underpinning considerations about *what* to measure, is *how* to measure it. In RCTs, surveys comprised of multiple-choice closed questions are often used. Undoubtedly, multiple choice

questions cannot provide detailed information about students' knowledge, understanding, problem solving and critical thinking skills in relation to the Holocaust (Bischoping, 1998; Chapman & Hale, 2017). The same concern is applicable to attitudinal measures, where complex beliefs and thought processes are often distilled to a selection of statements where respondents indicate how far they agree or disagree with them. However, when research questions seek to determine the proportion of people in large samples who have particular knowledge or hold particular attitudes, before and after participating in an education programme, then multiple choice questions are useful. Qualitative data is then critical to enrich the quantitative data and answer research questions about how a programme has been implemented, why it did or did not have an impact and how far students' conceptual understanding has developed. Consequently, qualitative research should form an important element of any RCT exploring the impact of Holocaust education.

Instruments used to measure attitudes may not detect changes because assumptions about the nature and scale of the impact of Holocaust education have been misdirected. For example, before learning about the Holocaust most students will have some understanding of the ramifications of prejudice and racism. As pointed out by Salmons (2010), racism is not wrong because of what happened at Treblinka; instead racism is wrong because it is an erroneous view which continues to cause widespread injustice and suffering. Most students will recognise this long before they learn about the Holocaust. Gordon et al's (2004) study is apposite because no significant attitude differences were found from before to after learning about the Holocaust and the authors suggested students' pre-existing low levels of antisemitism and high levels of tolerance could account for these findings. This issue was also candidly expressed by a student in Foster et al's study (2016, p. 86): 'I didn't stop being racist because of the Holocaust; I wasn't racist and then not racist. I've always not been racist'.

When reflecting on the impact of Holocaust education, knowledge is not always positioned as the outcome variable, but instead as the predictor variable. In other words, greater knowledge is hypothesised to lead to more tolerant attitudes. However, given the findings of Foster et al.'s (2016) study showing how prevalent students' misconceptions about the Holocaust were (even after learning about the subject in school), knowledge should be framed as an outcome variable. As discussed, it is vital that students can draw on sound historical knowledge in order to understand the Holocaust in a meaningful way (Feinberg & Totten, 2016; Foster et al., 2016). Additionally, if knowledge is found to predict attitude change then without more attention paid to what information students have learned (as well as what they have *not* learned and/or muddled up) the mechanisms that helped and/or hindered this attitudinal change will be unknown. To give one illustration: Foster et al., (2016) found the majority of students overestimated the proportion of people who were Jewish in the pre-war German population. In light of predictions that Holocaust education impacts upon students' attitudes towards prejudice and tolerance, it is imperative to explore if and how these attitudes are affected by not knowing the Jewish population was a minority group constituting less than 1% of the pre-war German population. In a nutshell, incomplete knowledge of the Holocaust will have *an impact on the impact* of Holocaust education. Consequently, the nature and extent of students' knowledge (and knowledge change) must be explored as part of any impact study.

These issues highlight the challenges of identifying, conceptualising and measuring the impact of learning about the Holocaust. It would be erroneous to argue that different knowledge, understanding, attitudes and behaviours have to be measurable and quantifiable in order to be 'real' and to have occurred. As Biesta (2015) cautioned, it is essential to avoid valuing only what can be measured. However, given the ubiquitous 'sense' that Holocaust education has an impact on young people and the time and funding invested in these

programmes, it is essential for the field to reflect on these issues, seek to develop systematic research, and be more nuanced in the conclusions made about impact.

Some considerations for future impact research in Holocaust education

Many educators will attest to the impact that learning about the Holocaust has on students, yet pinpointing and measuring this impact is a complex issue. Arguably, if learning about the Holocaust has some form of positive influence on students, then seeking to define and measure this impact is superfluous. However, having recognised that teachers and educators want students to learn *something* as a consequence of studying the Holocaust, those involved in Holocaust education urgently need to decide, define and measure what this *something* is. This is particularly important because of the unintended consequences that can arise from learning about a subject which is not only complex and challenging, but is often coloured by pre-existing notions from cultural and political representations of the Holocaust ubiquitous within contemporary society. A good case in point is the prevalent belief amongst students (including those who have learned about the Holocaust in school) that Holocaust perpetrators killed Jews because they feared for their own lives (see Foster et al., 2016). This narrow understanding will impact on the nature of students' attitudes about the perpetrators and how they apply this to events in contemporary society. In contrast, different attitudes and reflections will develop when students have informed knowledge and understanding that the perpetrators' inclination to torture and kill Jews was fostered through Nazi campaigns to increase widespread antisemitism, scurrilous claims by the Nazis about the Jewish threat, and the Nazis' actions to dehumanise all Jews (see Hayes, 2017).

Clearly, an RCT is not the only way to judge how Holocaust education has impacted on students' knowledge and/or attitudes. Teachers have made effective use of various assessment methods to determine students' progress long before researchers and policy

makers were espousing RCTs in schools. Thus, RCTs should be supplementary to teachers' usual formative and summative assessment processes. Nevertheless, specialised Holocaust education programmes are delivered to young people across the world, at considerable cost, with good intentions and high hopes the programmes will have an 'impact' on them. Yet, the nature and scale of this impact remains unknown, and therefore, large scale and methodologically robust studies are essential.

Evidence shows that students often gather information about the Holocaust through television, films, books and the Internet (Foster et al., 2016). Thus, certain knowledge and/or attitudes are likely to be in gestation before formal learning about the Holocaust takes place in schools, with students starting to build a personal awareness of and response to the Holocaust informed by sources varying in content, perspective and accuracy. In this way, the impact of Holocaust education delivered in schools will be influenced by what students already think about the Holocaust as well as how they interpret new information.

Consequently, it is important to explore students' 'starting points' as some will have greater knowledge of the Holocaust than others. Random allocation of classes to intervention and control groups can help to balance out differences across conditions. However, control groups and intervention groups will never be perfectly matched (Keppel, Saufley, & Tokunaga, 2002). This is something that those working in Holocaust education will have to recognise and assimilate in their interpretations of the results.

In experimental research, the control group could be students who have not yet learned about the Holocaust in school, or students who are learning about the Holocaust but whose teachers are not implementing the education programme of interest. In the latter scenario the nature and quality of Holocaust education across the classes/schools is likely to vary. Teachers in control groups may also change their teaching practice as a consequence of the study especially if they are worried their approach might prove less effective than the intervention

(Walser, 2014). With an increasing climate of accountability in schools teachers may feel like their practice is being judged. This could contribute to a situation where only teachers who are confident their students will do well will volunteer to participate (a particular problem in quasi-experiments and a source of additional bias in the data). Collectively, these issues point to the importance of carefully defining the nature of the control group, collecting qualitative data on the programme they receive, and drawing on this information when interpreting the data.

A model common in England is for specialist educators at Holocaust education organisations to provide professional development programmes (CPD) for teachers to learn about Holocaust education pedagogy, and then apply this to their own teaching practice. Thus, there is a risk that some elements of an education programme will become distorted and misremembered as they pass along the specialist educator – teacher – student chain. Indeed, Barr et al's (2015) study suggested that this CPD model has greater impact on teacher outcomes than student outcomes. Additionally, evidence has suggested that most education programmes are adapted by teachers for their own contexts which will have implications for the impact. These adaptations have the potential to have a detrimental or beneficial influence on the programme's impact, but either way they are interacting with outcomes and can contribute to a misleading picture of the programme's effectiveness (Lendrum & Humphrey, 2012). Consequently, researchers need to conduct fidelity analysis to determine the extent to which the programme was delivered as intended (Hanley, Chambers, & Haslam, 2016).

Identifying the teaching aims used in the programme of interest is critical because they will shape what students learn, and thus, will inform how 'impact' is operationalised and measured. If teaching aims tap into attitudes or behaviours, teaching content must address these aims (rather than assuming an 'osmosis' process has occurred to achieve these aims). Additionally, the complexity inherent in the measurement of attitude and behaviour change

needs to be recognised and additional or alternative concepts considered like the contribution of Holocaust education to the formation of different types of beliefs (drawing on Fishbein and Ajzen (2010) for example). Researchers also need to concede that Holocaust education may not be the principal predictor in the formation of particular attitudes, beliefs or behaviours and consider other variables which could interact with Holocaust education. Furthermore, it is essential that students' knowledge of the Holocaust is studied as an outcome variable. What students learn as a consequence of participating in a particular Holocaust education programme, and the richness and accuracy of their knowledge and understanding, will undoubtedly interact with their beliefs and attitudes.

Biesta (2007) reminds us, education is not a simplistic linear process between the teacher and student. Instead, the connection between teaching and learning emerges and evolves through students' interpretations of their educational encounters. This is something all researchers and educators who are exploring the impact of Holocaust education must keep in mind, especially given how complex and challenging learning about the Holocaust is on both a cognitive and affective level.

Conclusion

Despite the complexities of exploring the impact of Holocaust education, research in this area is essential if it is to provide insights into how learning about the subject affects young people – not only in the anticipated or hoped for ways (e.g. developing attitudes of acceptance and empathy), but also in unforeseen and unintended ways (e.g. inadvertently reinforcing pre-existing myths and misconceptions about the Holocaust leading to superficial or tainted attitudes). In England, the Holocaust continues to be a mandatory part of the National Curriculum with successive Westminster governments pledging their support (and funding) to Holocaust education and memorialisation. Thus, questions about the impact of Holocaust

education are likely to increase and evolve, especially given the greater utilisation of RCTs in education. However, RCTs require specialist expertise, research funding and require buy-in from teachers, educators and policy makers. Moreover, as identified in this paper, the complexity of Holocaust education – both in terms of the content delivered and how students respond to and assimilate this material – would make implementation and process evaluations (IPE) a necessary component of this work.

While it may appear this paper presents a bleak picture of the current state of impact research in Holocaust education, as a field we need to admit that research conducted to date has only provided limited and potentially fallible information about the impact of Holocaust education. It is imperative that as we move forward, the methods used are sound and systematic, and they enable educators, researchers and policy makers to talk about impact based on firm evidence rather than intuition.

Declaration of interest statement

No potential conflict of interest was reported by the author

Biographical note: Dr Rebecca Hale is a Senior Research Associate at the UCL Centre for Holocaust Education, part of the UCL Institute of Education. She began her career as a secondary school teacher before completing a Masters and PhD in psychology. She joined the UCL Centre for Holocaust Education in 2013 and has led a number of the Centre's projects, principally those with quantitative elements. Becky is one of the lead authors of *What do students know and understand about the Holocaust? Evidence from English Secondary Schools* and has published papers in the fields of psychology and education.

References

- Barr, D. J., Boulay, B., Selman, R. L., McCormick, R., Lowenstein, E., Gamse, B., ... Leonard, M. B. (2015). A randomized controlled trial of professional development for interdisciplinary civic education: Impacts on humanities teachers and their students. *Teachers College Record, 117*, 1-52.
- Bauer, Y. (2016). Foreword by Professor Yehuda Bauer. In S. Foster, A. Pettigrew, A. Pearce, R. Hale, A. Burgess, P. Salmons, & R-A. Lenga. *What do students know and understand about the Holocaust? Evidence from English secondary schools* (pp. viii-ix). London: UCL.
- Biesta, G. (2007). Why 'what works' won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory, 57*, 1 – 22.
- Biesta, G. (2015). What is education for? On good education, teacher judgement, and educational professionalism. *European Journal of Education, Research, Development and Policy, 50*, 75-87.
- Bischoping, K. (1998). Method and meaning in Holocaust-knowledge surveys. *Holocaust and Genocide Studies 12*, 454-474.
- Chapman, A., & Hale, R. (2017). Understanding what young people know: methodological and theoretical challenges in researching young people's knowledge and understanding of the Holocaust. *Holocaust Studies, 23*, 289-313.
- Coe, R. (2009). Unobserved but not unimportant: the effects of unmeasured variables on causal attributions. *Effective Education, 1*, 101-122.

Concato, J., Shah, N., & Horwitz, R.L. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *The New England Journal of Medicine*, 342, 1887-1892.

Connolly, P., Biggart, A., Miller, S., O'Hare, L., & Thurston, A. (2017). *Using randomised controlled trials in education*. London: Sage.

Cook, T. D. (2002). Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them. *Educational Evaluation and Policy Analysis*, 24, 175-199.

Cowan, P., & Maitles, H. (2007). Does addressing prejudice and discrimination through Holocaust education produce better citizens? *Educational Review*, 59, 115-130.

Eadson, W., Pearson, S., Foden, M., Sanderson, E., & Bashir, N. (2015). *Holocaust Memorial Day impact study: final report*. Sheffield Hallam University: Centre for Regional Economic and Social Research (CRESR). Retrieved from: <https://www.hmd.org.uk/wp-content/uploads/2018/06/HMD-Impact-Study-Final-Report-November-2015-PDF.pdf>

Education Endowment Foundation (2018). *Annual Report 2018*. London: Education Endowment Foundation. Retrieved from: https://educationendowmentfoundation.org.uk/public/files/Annual_Reports/EEF_-_2018_Annual_Report.pdf

Feinberg, S., & Totten, S. (2016). Foundational concerns: developing historically accurate and pedagogically sound Holocaust lessons and units. In S. Totten & S. Feinberg (Eds.), *Essentials of Holocaust education: Fundamental issues and approaches* (pp. 1-16). New York, NY: Routledge.

Fishbein, M. & Ajzen, I. (2010) *Predicting and changing behavior: The reasoned action approach*. New York, NY: Psychology Press.

Foster, S., Pettigrew, A., Pearce, A., Hale, R., Burgess, A., Salmons, P., & Lenga, R-A. (2016). *What do students know and understand about the Holocaust? Evidence from English secondary schools*. London: UCL.

Gorard, S. (2010). All evidence is equal: the flaw in statistical reasoning. *Oxford Review of Education*, 36, 63-77.

Gordon, S. B., Simon, C. A., & Weinberg L. (2004). The effects of Holocaust education on students' level of anti-semitism. *Educational Research Quarterly*, 27, 58-71.

Gross, M.H. (2017). Learning from survivors: mapping the past onto the present. *Intercultural Education*, 28, 591-606.

GOV.UK (2018, September 4). Public consultation marks next stage in development of UK Holocaust Memorial and Learning Centre. Retrieved from <https://www.gov.uk/government/news/public-consultation-marks-next-stage-in-development-of-uk-holocaust-memorial-and-learning-centre>.

GOV.UK (2019, May 7) Prime Minister leads unprecedented support for Holocaust Memorial as further £25m committed. Retrieved from <https://www.gov.uk/government/news/prime-minister-leads-unprecedented-support-for-holocaust-memorial-as-further-25m-committed>

Handley, M. A., Lyles, C. R., McCulloch, C., & Cattamanchi, A. (2018). Selecting and Improving Quasi-Experimental Designs in Effectiveness and Implementation Research. *Annual Review of Public Health*, 39, 5-25.

Hanley, P., Chambers, B., & Haslam, J. (2016). Reassessing RCTs as the 'gold standard': synergy not separatism in evaluation designs. *International Journal of Research & Method in Education*, 39, 287-298.

Hayes, P. (2017). *Why? Explaining the Holocaust*. New York, NY: W.W. Norton and Company, Inc.

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R., & Kerr, K. (2016). *Implementation and process evaluations (IPE) for interventions in education settings: An introductory handbook*. University of Manchester: Education Endowment Foundation.

Retrieved from:

https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Handbook.pdf

Jedwab, J. (2010). Measuring Holocaust knowledge and its impact: A Canadian case study. *Prospects*, 40, 273-287.

Karn, A. (2012). Toward a philosophy of Holocaust education: Teaching values without imposing agendas. *The History Teacher*, 45, 221-240.

Keppel, G., Saufley Jr, W. H., & Tokunaga, H. (2002). *Introduction to design and analysis: A student's handbook* (2nd ed.). New York, NY: WH Freeman and Company.

Kinloch, N. (1998). Learning about the Holocaust: moral or historical question? *Teaching History*, 93, 44-46.

Lendrum, A. & Humphrey, N. (2012). The importance of studying the implementation of interventions in school settings. *Oxford Review of Education*, 38, 635-652.

- Levy, S. A. & Sheppard, M. (2018). "Difficult knowledge" and the Holocaust in history education. In S. A. Metzger & L. M. Harris (Eds.), *The Wiley international handbook of history teaching and learning* (pp. 365-387). New Jersey, NJ: John Wiley & Sons, Inc.
- Lindquist, D. H. (2008). Five perspectives for teaching the Holocaust. *American Secondary Education*, 36, 4-14.
- Maxwell, J. A. (2004). Using qualitative methods for causal explanation. *Field Methods*, 16, 243-264.
- Maxwell, J. A. (2012). The importance of qualitative research for causal explanation in education. *Qualitative Enquiry*, 18, 655-661.
- National Audit Office (2018, February 22). Converting maintained schools to academies. Retrieved from <https://www.nao.org.uk/press-release/converting-maintained-schools-to-academies/>
- Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *BMJ*, 332, 413-416.
- Pearce, A. (2017). The Holocaust in the National Curriculum after 25 years. *Holocaust Studies*, 23, 231-262.
- Pettigrew, A. (2010). Limited lessons from the Holocaust? *Teaching History*, 141, 50-55.
- Pettigrew, A. (2017). Why teach or learn about the Holocaust? Teaching aims and student knowledge in English secondary schools. *Holocaust Studies*, 23, 263-288.
- Pettigrew, A., Foster, S., Howson, J., Salmons, P., Lenga, R-A. & Andrews, K. (2009). *Teaching about the Holocaust in English secondary schools: An empirical study of national trends, perspectives and practice*. London: Institute of Education.

Salmons, P. (2010). Universal meaning or historical understanding? The Holocaust in history and history in the curriculum. *Teaching History*, 141, 57-63.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31, 15-21.

Starratt, G. K., Fredotovic, I., Goodletty, S., & Starratt, C. (2017). Holocaust knowledge and Holocaust education experiences predict citizenship values among US adults. *Journal of Moral Education*, 46, 177-194.

Sullivan, G. M. (2011). Getting off the “gold standard”: Randomized controlled trials and education research. *Journal of Graduate Medical Education*, 3, 285-289.

Torgerson, D. J. & Torgerson, C. J. (2008). *Designing randomised trials in health, education and the social sciences*. Basingstoke: Palgrave Macmillan UK.

Whitty, G., & Wisby, E. (2016). Education(al) research and education policy in an imperfect world. In G. Whitty (Ed.) *Research and Policy in Education* (pp. 1-19). London: UCL Institute of Education Press.

UNESCO. (2014). *The impact of Holocaust education: how to assess policies and practice?* Paris: United Nations Educational, Scientific and Cultural Organization.

van Driel, B. (2003). Some reflections on the connection between Holocaust education and intercultural education. *Intercultural Education*, 14, 125-137.

Walser, T.M., (2014). Quasi-experiments in schools: the case for historical cohort control groups. *Practical Assessment, Research and Evaluation*, 19, 1-8.

Word count: 7,891

ⁱ Although Holocaust Education is mandatory for Key Stage 3 history in England, academies (funded directly by the government and run by an academy trust) do not have to follow the National Curriculum. In 2018, 72% of

secondary schools were academies (National Audit Office, 2018). Many academies implement the National Curriculum but given this is not compulsory, it is likely some schools will have opted out of teaching about the Holocaust (Foster et al., 2016).

ⁱⁱ Facing History and Ourselves is a non-profit teacher professional development organisation providing support to secondary school teachers and students to make connections between social history, norms of human behaviour, and the civic and moral choices confronted in their daily lives (see Barr et al., 2015, pp. 4-6).