

Amanda Spielman
Her Majesty's Chief Inspector
OFSTED
Clive House
70 Petty France
London
SW1H 9EX



20 August 2021

Dear Amanda,

Re: Ofsted's response to AMET's complaint about the review of mathematics research

My thanks go to Ofsted for making time to construct and publish openly the inspectorate's response to the concerns raised by AMET about the mathematics research review. Not only does it provide additional clarity on Ofsted's approach to reviewing extant research of quality in mathematics education, it is also a useful example of the inspectorate's way of dealing with feedback from peers on those reviewing activities. I would therefore like to take this opportunity to share some of my observations of the response in hope that Ofsted and members of the education community find them helpful.

AMET identified concerns with the way a significant proportion of citations were used in the review to support claims or statements made by Ofsted. From the association's analysis of the review, a 46-page table was produced which set out clearly where they found, and did not find, a substantive match between claims made in the review and the accompanying references. Details were also included where authors expressed concerns about the use of their research work. In addition, AMET raised an equally important but smaller number of concerns about other aspects of research practice and accompanied these with comprehensive comments alongside specific examples. Detailed comments were also included to support concerns about some methodological features of the review.

In Ofsted's response most of the issues raised by AMET are dismissed outright, although one is dismissed with a caveat, and some are not responded to at all. Given that Ofsted acknowledge that education research is contestable and contested, and hence has shared the reviews in order to obtain feedback from the broader subject community, it seems pertinent to consider the extent to which a worthy response to AMET's concerns has been provided by Ofsted.

In what follows I comment initially on Ofsted's response to concerns about the review methodology before moving on to Ofsted's response to concerns about the use of citations and other elements of research practice.

I refrain from including references to material contained in the mathematics review, with the exception of briefly mentioning some additional information about two of Ofsted's statements, for the purpose of clarification. In the interest of brevity, I also do not draw on extant research that relates to the comments I make, although there are plenty of studies available to readers should they wish to consult them. I do however make very occasional reference to Ofsted's paper 'Principles behind Ofsted's research reviews and subject reports' as this document outlines the protocol for the reviews.

The drawing together of research material has been a feature of education research for centuries. This has enabled a wide range of approaches for reviewing extant research to be developed, tested, shared and refined to meet the specific challenges of an extensive and diverse array of education settings. Each reviewing methodology has both a distinctive conceptual framework and operating protocol which are synchronised with the purpose of the study.

Noticeably, Ofsted has not chosen to deploy a contemporary method of reviewing research for their subject reviews. Instead, the inspectorate has developed its own distinctive approach to reviewing research material. An appraisal of the Ofsted methodology therefore needs to take account of the purpose of the review.

Ofsted have publicly stated that the form and function of the reviews are outlined in their research principles paper which was published on 30 March 2021. A specific aim of the reviews is to set out the research literature that has informed Ofsted's thinking about subject quality. Ofsted have used several criteria to act as filters when compiling the reviews to ensure that all of the literature included aligns with their principles for quality of education as outlined in the 'Education inspection framework (EIF)' and summarised in 'Education inspection framework: overview of research'.

It can therefore be argued, that if Ofsted had followed a traditional literature review approach and included research material that did not align with their principles for quality of education it would have been in breach of the published conceptual framework and protocol for the reviews. Consequently, on this footing Ofsted can justifiably defend the overall approach of the mathematics research review because basically it appears to have been carried out in the way it was designed to be carried out.

Understandably, this approach has led to the production of a document that is essentially a declaration of Ofsted's thinking about subject quality. For this reason, it should not be viewed as a conventional review of research on high-quality school mathematics education.

AMET share the view that Ofsted's approach is unlike a conventional literature review which usually develops a thesis from considering at times contrasting viewpoints in the literature consulted whereas Ofsted's approach does in effect start with a thesis and then searches for literature to support it. With this in mind AMET question the aptness of the title of Ofsted's material and suggest that it would be clearer to the education community to call these documents position papers rather than research reviews.

Ofsted dismiss this concern and defend the use of the term research review by claiming that the term is suitable because it is 'not generally used in any fixed way'. This is a controversial claim to make considering the rich history of reviewing practice in education, which may be why Ofsted do not furnish any evidence or an argument to support it.

Ofsted also state clearly that the process of putting together the mathematics review included theorisation of the conception of quality of mathematics education that emerged from the findings in the literature accessed and the intention was not to write a conventional literature review. So, Ofsted's unconventional reviewing methodology must not be mistaken for that used in a more conventional review of research on subject quality.

To avoid confusion, it would thus have been much better to have given the documents in the series titles that reflected more closely the aim of producing declarations of Ofsted's thinking about subject quality. The existing titles make it too easy to confuse them with reviews of research on high-quality school subject education that have followed conventional reviewing methodologies. As a consequence, the potential of the documents to mislead readers in this respect is very high.

In addition to AMET's concern above, the association pointed out that the research presented also appears to be uncontested because only supporting sources have been used. The result is that the research base appears to be stronger than it is.

Ofsted dismiss this concern and defend the inclusion of only supportive literature by referring to their principles paper and their review of research for the education inspection framework. In addition, Ofsted point out that within the parameters outlined in those documents the process of the review did include the consideration of sources with contrasting viewpoints, but no evidence is provided to support this

assertion. Ofsted also clarify that the research review format does not outline these deliberations because including this explanation would make the research review less accessible.

As mentioned previously Ofsted can choose to justify the exclusion of material in the review that contradicts or contrasts with their principles for quality of education because this is the intention of the inspectorate's approach to reviewing the literature. But AMET's concern is about the effect of using this strategy and how, in particular, it has created the impression of providing a base of research that is stronger than it actually is. Ofsted's response does not therefore address directly AMET's concern.

Interestingly, Ofsted mention that the process of the review did include the consideration of sources with contrasting viewpoints. If that was the case, this would have been in breach of the research protocol which includes filters to ensure that all of the literature included aligns with their principles for quality of education, unless of course by consideration Ofsted simply mean the use of a basic accept/reject decision framework for the literature accessed.

It is not clear from Ofsted's response what accessibility means in relation to the review or indeed how the inclusion of an outline of deliberations about sources of contrasting viewpoints would have made the review less accessible. Most research reviews include discussions of contradictory and/or contrasting findings but that doesn't necessarily reduce the accessibility of the material.

AMET rightly pointed out that it is common practice for reviews of research to highlight some key details about a study being cited to help readers to judge the strength of the claim it is being used to support. The association noted that this was not done in this review and it is only by tracking down and reading the individual studies that this can be determined.

Ofsted does not provide a response to this concern.

In all, AMET identified concerns with the way 165 citations were used to support claims or statements made by Ofsted. This is more than half (54%) of the 307 used in the review. AMET have rightly drawn attention to the fact that when people read a paper, they need to be able to trust that the references cited support the point that is being made.

This is a particularly important point to mention as we currently inhabit an education world that has never witnessed the production and dissemination of education research material at today's speed and volume. Consequently, teachers, leaders, researchers and policy makers often find it difficult to keep abreast of things as they develop. Understandably, where they believe they can, they turn to and rely upon others to help them out. When those to whom they turn have significant influence in the education community they are thus inclined to uncritically accept what is offered. The concerns AMET identified with Ofsted's citation and referencing practice therefore warrant detailed and comprehensive attention.

Ofsted dismiss all of these concerns and defend the way references have been used in the review. In order to create a response to these concerns Ofsted purport to have sampled 35 papers from the index of the 165 citations identified but do not provide evidence of doing so. Ofsted claim to have not found any examples where the references did not support the points made by Ofsted but do not provide evidence to support this claim. Ofsted note that it is possible that the concerns may have arisen as a result of some misinterpretation of the form and function of the mathematics research review and this may have been avoided if further information had been provided. Ofsted go on to provide notes on interpretation for 5 specific sources (3%) of the 165 queries raised by AMET. Ofsted uses the reported outcome of the sample analysis and the notes on source interpretation to defend the stance that the claims made in the review are supported properly by the references used.

Now, given the critical importance of ensuring claims are supported by substantive evidence in any piece of research it is not unreasonable to query why only 35 out of 165 instances may have been investigated by Ofsted, especially when AMET had gone to the trouble of expending significant time and expertise to

review all of the 307 citations in the review as part of the association's analysis. It is also not unreasonable to seek more detail on how the sample of 35 was selected and most importantly how the five specific sources were selected too. Whilst I shall not comment on the credibility of the notes on interpretation here, (I believe members of the subject community are very likely to do that for themselves) it clearly isn't realistic to infer that all citations match the claims made in the review based on this methodology. Evidence relating to 3% of the concerns raised, no matter how credible that evidence may be, is not sufficient to draw a robust conclusion about all of the concerns let alone about all of the citations.

To illustrate my point further let's say I inform readers that I selected 35 references from those included in the 165 cited and state that I DID NOT find any that contradicted the concerns raised by AMET. This assertion alone provides very little evidence to support the view that AMET's concerns are valid because readers have no other option than to take my word for the assertion.

By way of adding evidence to support my assertion let's now consider my interpretation of one of the studies from the sample I selected: Footnote 86. J-W Son and S Senk, 'How reform curricula in the USA and Korea present multiplication and division of fractions', in 'Educational Studies in Mathematics', Volume 74, Issue 2, 2010, pages 117 to 142.

In the review Ofsted used this material to support claims relating to the use of informal methods when teaching mathematics and specifically to using them for only a short amount of time, as a bridge to formal written methods.

AMET made the following comment about this research in relation to the claim it is used to support: 'Not a match: informal methods not discussed; the focus was on conceptual and procedural understanding and the advantage of doing these together'.

Having reviewed the paper I decide to quote the following from the abstract of the study: 'In order to give insights into cross-national differences in schooling, this study analysed the development of multiplication and division of fractions in two curricula: Everyday Mathematics (EM) from the USA and the 7th Korean mathematics curriculum (KM)..... Both curricula provide opportunities to develop conceptual understanding and procedural fluency. However, in EM, conceptual understanding is developed first followed by procedural fluency, whereas in KM, they are developed simultaneously. The majority of fraction multiplication and division problems in both curricula requires only procedural knowledge. However, multistep computational problems are more common in KM than in EM, and the response types are also more varied in KM.'

This extract clearly confirms that the focus of the study was on conceptual and procedural understanding. Based on the evidence I quote from my review of Son and Senk's study, I am now able to say with a high degree of confidence that AMET's point about the focus of the study is valid.

If I mention in addition that when I read the paper I noticed there was no discussion of informal methods and when I used a search facility on the study to find references to informal methods it yielded a null return, I am also in a position to say with a high degree of confidence that AMET's point about informal methods NOT being discussed in the study is valid.

When this point is coupled with the evidence of the focus of the study it is fair and reasonable to state that the concern raised by AMET about the use of this reference is upheld and that the reference has been used inappropriately to support the claims made by Ofsted in the review.

However, reporting one example that supports AMET's concerns about the use of references, alongside the simple statement that I did not find any instances of the use of references that contradicted the concerns raised by AMET, cannot be justifiably used to claim with sufficient confidence that all of the remaining 164 concerns are valid. What's more, there is little point in me providing a total of 5 examples in support of my claim because even if all of them enabled me to say with the same high degree of

confidence that AMET's concerns about the use of those references were valid, I would still know nothing at all about the remaining 160 concerns. Knowing something about only 5 out of 165 concerns whilst knowing nothing about 160 of them is not a secure platform for making a judgement about all of them.

In the same vein, the evidence and arguments Ofsted has provided are thus not sufficient to successfully dismiss AMET's concerns about poor citation practice in the mathematics review.

But Ofsted aren't just dismissing the concerns raised by AMET. Ofsted are also defending the use of references throughout the review. Ofsted are in effect using evidence to support the effective use of 5 citations, however robust that evidence may be, as a means of drawing a conclusion about the use of 307 citations. It is patently clear that Ofsted have also provided insufficient evidence and argument to successfully defend a claim of effective use of citations throughout the review.

It is worth noting of course that the points above do not render the use of Son and Senk's material inappropriate in the review rather, as with all the other literature accessed, Ofsted would have made more appropriate use of this material had its use been restricted to discussions that reflect better the intentions of the researchers and the context of the study.

The inclusion of a wide range of types of research in the review was welcomed by AMET. The association rightly pointed out that care needs to be taken when using findings from some types of research to support generalisations and by way of example attention was drawn to the use of a single study that involved research with four children in the United States to make a generalisation about task difficulty influencing the likelihood of pupils engaging in disruptive behaviour.

Whilst acknowledging that the study alone provided less robust evidence than multiple studies would have done, Ofsted defends the use of the single study to support the claim made by referring to the weight of evidence provided earlier in the review and to the studies referenced in the US paper. The logic of this response is somewhat puzzling. If the weight of evidence situated previously in the review is sufficient to support the claim made, there appears to be no good reason to have cited the paper at that point in the review. Alternatively, if the best evidence to support the claim is located in the references used in the US paper, then the appropriate references should have been cited instead of the US paper itself. Either and both ways, as it stands, the argument given for supporting the use of the paper in relation to the claim is in effect more of an argument for its inappropriate use than one for it being used suitably.

The inclusion of research from a wide range of countries in the review was also welcomed by AMET. The association rightly pointed out that care needs to be taken when using findings from research carried out in different countries not only because education has a complex relationship with culture, but also because different countries have different education systems that can perform quite differently especially on international comparisons. With this in mind concern was expressed about the balance of international research in relation to English/UK studies in the review. Attention was drawn specifically to the exclusion of a rationale for including a relatively high volume of research from the USA.

Ofsted dismiss this concern and defend the use of the balance of international research in the review by pointing out that the search strategy was not limited to any country as set out in the principles paper and that the inspectorate was 'limited by what we could access in English, but overall, we used research from a range of countries.' The response goes on to point out that 'where research in the US evidences effective practice, for example that concerned with the attainment and progress of disadvantaged children and children with SEND, that does not mean it is put into practice widely in US schools, from which pupils are drawn for international comparisons, such as the OECD's PISA tests.'

This response is largely unrelated to AMET's concern. AMET complimented Ofsted on the use of a search strategy which embraces research from a range of countries. There is no need for Ofsted to make a point of confirming this approach. The inclusion of findings from research of effective practice from the US,

irrespective of the context of the research, would clearly be included in the review if it was accessed and aligned with Ofsted's principles for quality of education.

Furthermore, the response does not address AMET's concerns about the balance of international research in relation to English/UK studies nor does it provide reasons for the inclusion of a relatively high volume of research from the USA other than to mention that the inspectorate was 'limited by what we could access in English' which may indicate problems with the search strategy and/or highlight doubts about Ofsted's resources to carry out the research.

The age of some of the sources was a concern for AMET. The association noted that the review includes research ranging from 1939 to the present day with 'approximately 50% of the references coming from sources published within the last 10 years.' AMET pointed out that nearly half of the sources were published before the current National Curriculum was implemented which can result in discussing practices that no longer occur. By way of an example attention was drawn to a paper from 1988 which was used to support a statement related to computer use.

Ofsted dismiss this concern and defend the age of the research material used in the review by asserting that the search strategy focused primarily on research conducted since 2010 and explaining that 'where there were seminal works or we could not find good quality literature on a specific research question, we went further back. Good quality being defined by our principles paper.' Noticeably, Ofsted do not furnish evidence to support this claim.

As mentioned previously Ofsted have repeatedly pointed out that the search strategy for the review was guided by the inspectorate's principles paper. It is interesting to note therefore that Ofsted claim that the primary focus of the search strategy was on research conducted since 2010, given there is no mention of adopting this approach in the principles paper.

Moreover, had Ofsted emphasised the search in this way it seems very unlikely that around 50% of the references used in the review would have come from sources outside of this focus. It is not unreasonable to expect a research strategy that focuses primarily on a specific age range of research publications to produce a majority of sources that fall within the specified age range. The inclusion of an almost 50:50 split between research material published since 2010 and that which is older does not reflect a primary focus on research conducted since 2010.

Even if as Ofsted suggest, in order to gather good quality literature, as defined by Ofsted, in relation to specific research questions it was necessary to access material published prior to 2010, it is also very unlikely that the deficiency in more contemporary suitable research would amount to almost 50% of that collected, even when allowance is made for the inclusion of seminal works, unless of course only a handful of research is available.

What's more, Ofsted do not address AMET's concern about the use of research material from 1988 to support a statement made about the use of computers.

In addition to AMET's earlier point about readers needing to track down individual studies to identify the key details of a study being cited, the association pointed out this needs to be done specifically to identify the age groups involved in the material presented. AMET noted that whilst many of the studies included in the review involved school children, most of these were conducted with a specific phase or year group and they have often been used to support statements that generalise beyond these ages. By way of example AMET pointed out the use of a study with 1st and 2nd grade children (6-8 years old) in the USA being erroneously equated to the 'start of the academic journey' in England, which would be nursery or reception classes (3-5 years old).

Ofsted's response to this concern is short, dismissing it and defending the frequent use of studies of specific age groups to make generalisations by simply asserting that 'where claims are made about phases

of education, alignment/overlap has been sought.' Ofsted do not provide evidence or argument to support this assertion and do not describe or explain what is meant by alignment/overlap in this context.

Instead, Ofsted choose to respond in some detail to the use of the USA study, confirming that it was a study of children in grades one to two in the US (mostly six- to seven-year-olds) and point out that it was used to support a statement about 'the 'start of the academic journey' as a period where 'maths anxiety' can manifest' and 'for some, the 'start of the academic journey' can be at any point in their lives. Ofsted go on to claim 'there is not much disparity between the source age group and the age group being referred to in the research review.' Ofsted do not furnish evidence or argument to support this claim and do not offer an explanation of what is meant by disparity in this context.

Now, even if Ofsted had supplied evidence to suggest that there is some commonality between the mathematics anxiety of 3-5 years old and 6-8 years old pupils, Ofsted would still have needed to provide evidence of using a credible means of establishing that this was sufficient to enable generalisations to be made to any school age group, given Ofsted's statement that the 'start of the academic journey' can be at any point in their lives.

In addition, Ofsted assert in relation to the claim 'This is because the research review states that the conception of quality outlined is applicable to reception year onwards (not nursery stages) and the fact that maths anxiety in US grade one is likely to have some of its origins in kindergarten, which is approximate to reception year and is, in many states, compulsory.'

When taking in to consideration the points above, Ofsted's response the AMET's concern about the inspectorate using studies conducted with a specific phase or year group to support statements that generalise beyond these ages is unclear, problematic and lacks sufficient evidence to support it.

To sum up. In this response I have examined Ofsted's response to the concerns AMET raised about the quality of the mathematics research review. I have commented on Ofsted's responses to both concerns about the methodology used and to concerns about specific elements of research practice, including the use of citations.

It is important to recognise that whilst I have made a point of noting that Ofsted is in a position to make a case to defend the approach used for reviewing the literature accessed, this must not be taken as support for the methodology itself. I have focussed only on Ofsted's response to the concerns raised by AMET. I have not provided a detailed analysis of Ofsted's approach.

Overall, I have found that the inspectorate has not provided sufficient evidence or argument to support the dismissal of the concerns raised by AMET or to support the defence of the stances taken in Ofsted's response. I have also found instances where Ofsted has not responded at all to AMET's concerns.

This leads me to the conclusion that Ofsted's response can be best described as inadequate, and as a consequence the concerns raised AMET remain standing.

Yours sincerely,

A handwritten signature in black ink, appearing to read 'Terry Pearson', with a stylized flourish at the end.

Terry Pearson
Independent researcher

Sent by email to: enquiries@ofsted.gov.uk

c.c. info@ametonline.org.uk (personal information redacted)

Posted as reply to: https://www.twitlonger.com/show/n_1srpfg6 (response comments only)