OpenDevEd
Development & Education

Protocol Addendum for a Systematic Review and Meta-analysis

Understanding the quality of EdTech interventions and implementation for disadvantaged pupils

Bethany Huntington

Christopher Klune

Aditi Bhutoria

Björn Haßler

October 2024

10.53832/opendeved.1161

https://opendeved.net

Education Endowment Foundation

The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus – Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

For more information about the EEF or this report, please contact:

Education Endowment Foundation

5th Floor, Millbank Tower

21–24 Millbank

SW1P 4QP

0207 802 1653

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

About Open Development & Education

Open Development & Education is an organisation established in 2014, committed to open and transparent approaches to development and education, including organisational, professional, and personal development and education.

We work with education institutions, governments, donors, aid agencies, and other stakeholders to conduct research and develop effective strategies and policies for furthering equity. Equity, evidence, and accessibility are at the heart of our approach. In our implementation and research, we seek to support and build local capacity. Our Evidence Library is an online repository of all our publications. It houses a wealth of the latest research on education, educational technology, and more.

To learn more about us and our work, visit https://opendeved.net. You can also explore our Evidence Library at https://docs.opendeved.net/lib/

Acknowledgements

We would like to thank Dr Louis Major for providing the research team with 'critical friend' feedback on drafts of the research protocol and subsequent addendum.

About this document

Recommended citation

Huntington, B., Klune, C., Bhutoria, A., & Haßler, B. (2024). *Protocol Addendum for a Systematic Review with Meta-Analysis: Understanding quality characteristics of Edtech interventions and implementation for disadvantaged pupils.* Open Development & Education. https://doi.org/10.53832/opendeved.1161. Available under Creative Commons Attribution 4.0.

Key project publications

Haßler, B., McBurnie, C., Walker, H., Klune, C., Huntington, B., & Bhutoria, A. (2024). *Protocol for a Systematic Review with Meta-Analysis: Understanding Quality Characteristics of Edtech Interventions and Implementation for Disadvantaged Pupils.* Open Development & Education. https://doi.org/10.53832/opendeved.1077. Available under Creative Commons Attribution 4.0 International https://creativecommons.org/licenses/by/4.0/deed.en

Contents

Deviations from the original protocol

Since the initial protocol was published (↑Haßler et al., 2024a), several deviations from the original methodological plan were made in response to practical considerations that arose during the research and data extraction processes. These changes, detailed in Table 1 below, were deemed necessary to maintain the highest quality of research while ensuring alignment with the study objectives. Each adjustment was carefully implemented to ensure the research remains systematic and rigorous throughout.

Table 1. Deviations from the methodology outlined in the initial protocol

Original specification	Deviation
The inclusion criteria in the protocol stated that for a comparison group, the following options would be included: ■ ***No intervention, non-EdTech intervention, or waitlist intervention for a control group*** ■ ***Another EdTech intervention as a comparison group*** These criteria were initially set to explore a broader range of comparisons, including how different types of EdTech interventions compared against each other and how they performed relative to both technology-based and non-technology-based controls.	After discussion between the research team and EEF when initially screening, we amended our inclusion criteria so that only the following groups were included in the meta-analysis: ***No intervention, non-EdTech intervention, or waitlist intervention for a control group*** This amendment meant that studies were only included if they had a non-technology-based control group. This change was made to better isolate the specific impacts of EdTech on student attainment. By focusing on studies with non-technology controls, we can more accurately assess the added value of EdTech over alternative teaching methods and may better identify the mechanisms facilitating improved attainment.

For data extraction, we committed to extracting data from the studies identified as 'high quality'. We said that we would extract from 'medium-quality' studies, where only limited high-quality studies were found.	Despite identifying sufficient high-quality studies, we extracted data from both high- and medium-quality studies. This decision was made based on the following considerations, and in agreement with the EEF: ■ The quality assessment tool we used had no predefined score thresholds for categorising studies as high- or medium-quality, which created a degree of subjectivity in assigning ratings. While multiple raters conducted quality appraisals, the differences between high- and medium-quality studies were often minimal. The team agreed that excluding medium-quality studies would risk omitting valuable data and potentially significant insights. ■ Adopting this strategy also helped address potential biases in the tool's assessment of quasi-experimental designs (QEDs), which may have been rated more stringently compared to randomised controlled trials (RCTs) and mixed-methods (MM) studies. This is because QEDs were evaluated on factors like representativeness, which is hard to assess / document accurately in research studies. In contrast, RCTs and MM studies had clearer, more flexible assessment criteria. Only studies meeting a baseline level of methodological rigour were included to ensure robustness, and the addition of medium-quality studies is not anticipated to impact the findings negatively. To check this is the case, a sensitivity analysis will be conducted to determine the impact that including medium-quality studies had on the results. This approach allowed for a more comprehensive analysis while maintaining the integrity of the results.

Original specification	Deviation
Several data extraction decisions were not stated in the protocol, but were made by the research team prior to the extraction process commencing.	**Primary outcome** Where a study had multiple attainment outcomes, we focused on literacy (specifically reading comprehension) and mathematics outcomes over other subjects (e.g., history, science). **Multiple measures for the same subject** In cases where there were multiple attainment measures for the same subject, we prioritised those more aligned with the EEF toolkit (e.g., reading comprehension) and as recommended by the EEF effect size data extraction (ESDE) guide, gave preference to standardised measures over those developed by researchers specifically for the study. **Age groups / school grades** In consultation with the EEF, we decided to combine outcome data across different grade levels to have one total effect size for each outcome measure. For example, if the study reported outcome data for Grades 1, 2, and 3, we aggregated the outcomes to represent the overall impact of the intervention on the combined cohort. The expectation was that where grades span both primary and secondary levels, the extraction was kept separate and not combined. This allowed for a distinct effect size for each primary study to reflect the differing educational contexts and learning stages.

Additional details for the analytical plan

The research team confirms that the following data analysis plan was created during the data extraction phase of the project, prior to any analysis beginning.

Heterogeneity

In the protocol, we indicated that the full analytical plan would be outlined once the heterogeneity had been calculated. However, we are publishing this addendum prior to the heterogeneity calculations.

The thresholds we will use to assess heterogeneity will be the following, as outlined in the protocol and recommended by the Cochrane Statistical Methods Group (↑Deeks et al., 2019):

I^2 statistic thresholds:

- 0–40%: might not be important
- 30–60%: may represent moderate heterogeneity
- 50–90%: may represent substantial heterogeneity
- 75–100%: considerable heterogeneity

The interpretation of the observed I^2 statistic depends on both the magnitude and direction of effects and the strength of evidence for heterogeneity.

While these thresholds are commonly used in educational research, it is widely recognised that there is often high heterogeneity in education research due to the diverse groups of learners, settings, and learning approaches that are investigated (↑Bernard et al., 2023; ↑Cheung & Slavin, 2012). As a result, we are prepared for elevated I^2 values.)

In cases of 'considerable heterogeneity' (75–100%), we will not report an overall mean effect size for that analysis, but will instead provide a forest plot to visualise the distribution of effect sizes (as seen in ↑Sims et al., 2021). Statistical heterogeneity will be transparently reported, and we will recommend that results be interpreted with caution due to the high heterogeneity between studies. Finding high heterogeneity will also not restrict the planned subgroup analyses that we plan to explore, as we anticipate that we will still gain important insights from the planned analyses despite potential conceptual / characteristic variations. Our approach will remain flexible, with heterogeneity providing context rather than being used as a barrier to analysis. By openly reporting heterogeneity levels

and structuring our analyses accordingly, we align with realistic thresholds in educational research while maintaining methodological rigour.

Subgroup analyses

Our minimum threshold for calculating a pooled effect size for the full meta-analysis is 30 studies. This aligns with existing practices in educational meta-analyses, where researchers have set a minimum threshold of 20–30 studies to ensure sufficient statistical power and reliability (↑Chaudhary & Singh, 2022; ↑Pigott & Polanin, 2020). We will require a minimum of 10 studies reporting relevant data for each potential subgroup analysis to ensure a sufficient sample size for robust analysis. Educational research frequently deals with diverse interventions and varying study designs, which can introduce significant variability. By establishing a smaller subgroup threshold of 10 studies, we can explore specific differences while maintaining the integrity of our findings (↑Cheung & Vijayakumar, 2016; ↑Koçak et al., 2021). This approach reflects a commitment to maintaining a rigorous methodology and is consistent with standards in educational research.

Depending on the availability of at least ten studies per subgroup, we will determine whether to conduct subgroup analyses for each of the following areas of interest. This will allow us to understand if variability in effect sizes across studies can be attributed to these factors:

- Disadvantage (e.g., Free School Meals (FSM) pupils, children in local authority care)
- School stage (e.g., elementary / primary, secondary, tertiary / post-secondary)
- Country (e.g., UK vs other countries with high technological readiness)
- Specific building blocks / mechanisms found to be prominent in the content analysis (e.g., automated feedback, personalised instruction, interactive app features – see Section 2.4 for more information on the content analysis)
- Subject domain (e.g., literacy, numeracy)
- Technological hardware (e.g., tablets, laptops)
- Technological software (e.g., adaptive learning platforms, e-reader software)
- Learning approach (e.g., blended learning, classroom learning, homework)
- Prominent thematic topics identified in the literature using Latent Dirichlet Allocation (LDA) topic modelling (e.g., personalised learning)

Implementation factors and participant characteristics may be explored through a moderator analysis, such as age, gender, delivery approach, intervention duration / frequency.

Intermediate outcomes

As part of this research, we will aim to investigate the following intermediate outcomes that may be associated with improved pupil attainment in EdTech interventions, as identified in the protocol:

- Student attitudes
- Student engagement
- Student motivation
- Student participation
- Student school attendance

During the data extraction process, the presence of intermediate outcomes was marked. Then, when a paper was confirmed for analysis, the intermediate outcome data was extracted alongside the primary attainment data.

We will analyse an intermediate outcome if at least ten studies report data for each given outcome (e.g., motivation). We will aggregate data from these studies measuring the same intermediate outcome to calculate an overall pooled effect size. A random-effects model will be used to account for variation between studies. These results will allow us to identify the overall effect of EdTech on these intermediate outcomes.

We will then conduct a Pearson's correlation analysis to explore the relationship between the intermediate outcomes and academic attainment. Pearson's correlation coefficient is suitable for this analysis as it measures the strength and direction of the linear relationship between two continuous variables – in this case, the effect sizes for intermediate outcomes and academic attainment (↑Ebenezer & Atakpa, 2023). This will allow us to identify whether higher intermediate outcomes (e.g., higher motivation) are associated with greater improvements in student achievement.

The coefficient will indicate the strength of the linear relationship, with values ranging from -1 (perfect negative correlation) to +1 (perfect positive correlation) and 0 indicating no correlation. We will interpret the strength of the correlation using common benchmarks within research guidance (↑Hinkle et al., 2003) (see below).

Size of Correlation	Interpretation
.90 to 1.00 (−.90 to −1.00)	Very high correlation
.70 to .90 (−.70 to −.90)	High correlation
.50 to .70 (−.50 to −.70)	Moderate correlation
.30 to .50 (−.30 to −.50)	Low correlation
.00 to .30 (.00 to −.30)	Negligible correlation

Mechanisms – Qualitative analysis

During the data extraction process, the research team identified and highlighted mechanisms within each study included in the meta-analysis and conducted initial semantic coding to summarise the mechanisms. A supplementary coding tool (↑Haßler et al., 2024b) was developed by Open Development & Education (OpenDevEd) and reviewed by the EEF. This tool provided the researchers with a systematic and rigorous protocol to extract and summarise mechanism data from the included literature. The research team was trained over two sessions by the research lead on identifying mechanisms in the literature and summarising them appropriately. Mechanisms were identified and summarised in accordance with the definition of mechanisms adopted from ↑Illari & Williamson (2012) in the research protocol (↑Haßler et al., 2024a). Mechanisms were identified in-text when researchers identified the three definitional components: the entity, the activity, and the outcome.

Once relevant text was highlighted, researchers summarised the identified mechanism, including these three components. These summaries constituted the extracted mechanism and coding. Studies reviewed by researchers may include several distinct mechanisms, while in some studies, no mechanisms may be identified.

The extracted mechanisms (highlighted text and summaries) will be imported into Atlas.ti to be qualitatively analysed by two researchers, including the research lead who helped develop the initial mechanism map developed in Phase 1 of the research and provided training on the mechanisms' data extraction. These initial

mechanisms from Phase 1 were structured in a sequential 'if… then… so' format outlined in the research protocol (↑Haßler et al., 2024a, p. 12) to capture the cause, initial effects, and outcomes within a distinct mechanism of effective EdTech use. This phase also iterated 12 initial building blocks through the inductive analysis of the initial mechanisms.

The continued development of the mechanism map over the project's three phases is a highly iterative process, informed by Phase 1 data and emergent data from ongoing research in Phase 2. As such, a hybrid deductive-inductive content analysis will be used to analyse extracted mechanisms data in refining the map and building blocks. Qualitative content analysis is a complex non-linear process, uniquely distinguished by the research goals and questions and how researchers adapt their methods accordingly (↑Elo & Kyngäs, 2008; ↑Mayring, 2000). As stated in the research protocol (↑Haßler et al., 2024a, p. 19), due to the iterative nature of the mechanism map, a combination of deductive and inductive analysis is needed to build on existing data and be responsive to emergent trends throughout the analysis. The approach is described below:

- The research team will create a coding framework based on the Phase 1 building blocks that will guide a content analysis in Atlas.ti. Using this framework, two researchers will conduct a preliminary round of coding on 20% of the extracted mechanisms to establish inter-coder reliability. Once this initial coding is complete, the coded data will be exported from Atlas.ti and into a statistical software package (e.g., SPSS/R). In this software, Cohen's Kappa will be calculated to determine the level of agreement between coders, with 0.70 considered an acceptable level of agreement (↑He et al., 2022).
- The two researchers will meet after the initial coding round to discuss and resolve any discrepancies. This may include refining the coding framework and adding new building blocks if deemed necessary to better encompass the coded mechanisms.
- If the initial agreement is acceptable, the researchers will continue coding independently until the full sample has been coded. If the initial agreement is less than 0.70, the coders will double-code a further 10% of the mechanisms. They will then reconvene to discuss any discrepancies and refine the coding scheme as needed to enhance clarity. If the agreement remains below 0.70, they will continue with iterative rounds of double-coding and discussion until they reach an acceptable level of agreement.

- After coding all the data into the framework, researchers will synthesise the codes within the building blocks in the framework. This will involve a review of the codes corresponding to each mechanism to see how prominent each building block is within the framework. In doing this, mechanisms will be refined and the building blocks revised. We will focus on how each mechanism is described and how it contributes to pupil attainment or intermediate outcomes (e.g., higher engagement, motivation, etc.). This will help us clarify how specific mechanisms interrelate.

    a. Any newly identified mechanisms from the data extraction (i.e., the highlighted text and summaries) will be re-described in the 'if… then… so' format in finalising the mechanism.

    b. Mechanisms from the initial Phase 1 map that are evidenced in the literature will be consolidated. This will allow us to identify the most prominent mechanisms shown throughout the literature and practitioner experiences, highlighting which mechanisms we can potentially explore quantitatively using subgroup analyses (in instances where ten papers report the same mechanism).

- Lastly, the building blocks from Phase 1 will be revised based on the newly identified and consolidated mechanisms. This may involve refining existing building blocks and / or creating new ones based on mechanisms that may not cleanly fit into existing ones. This stage allows us to systematically examine the predefined building blocks from Phase 1 and more rigorously iterate them to the mechanisms present in the primary research studies found during the meta-analysis.

In sum, analysing new mechanisms, consolidating existing ones, and revising the building blocks from the systematic review will ground the synthesis of a more robust and rigorous mechanism map. This will then be used as an artefact for the Phase 3 Structured Community Review, which will be further revised based on practitioner feedback and finalised (see ↑Haßler et al., 2024a, pp. 45–48). We will ask education practitioners to share evidence and practice-based insights to critique and validate conclusions from the Phase 2 analysis. We will validate our understanding of mechanisms and building blocks related to attainment, intermediate outcomes associated with attainment, linkages between mechanisms, and the relevance and applicability of specific mechanisms to disadvantaged pupils based on their classroom experiences.

References

This bibliography is available digitally in our evidence library at https://docs.opendeved.net/lib/IV8CRVJV

Bernard, R. M., Borokhovski, E., Schmid, R. F., & Tamim, R. M. (2023). Gauging the Effectiveness of Educational Technology Integration in Education: What the Best-Quality Meta-analyses Tell Us. In J. M. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy* (pp. 3929–3952). Springer International Publishing. https://doi.org/10.1007/978-3-319-17461-7_109. (details)

Chaudhary, P., & Singh, R. K. (2022). A meta analysis of factors affecting teaching and student learning in higher education. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.824504. (details)

Cheung, A. C. K., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review*, *7*(3), 198–215. https://doi.org/10.1016/j.edurev.2012.05.002. (details)

Cheung, M. W.-L., & Vijayakumar, R. (2016). A guide to conducting a meta-analysis. *Neuropsychology Review*, *26*(2), 121–128. https://doi.org/10.1007/s11065-016-9319-z. (details)

Deeks, J. J., Higgins, J. P., Altman, D. G., & on behalf of the Cochrane Statistical Methods Group. (2019). Analysing data and undertaking meta-analyses. In *Cochrane Handbook for Systematic Reviews of Interventions* (pp. 241–284). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119536604.ch10. (details)

Ebenezer, E., & Atakpa, A. O. (2023). Assessing the correlational statistics: The direction, weight and the relevance in educational research. *Intercontinental Journal of Education, Science and Technology*, *7*(1). (details)

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x. Available from (details)

Haßler, B., McBurnie, C., Walker, H., Klune, C., & Huntington, B. (2024). *Supplementary Coding and Data Extraction Tool for a Systematic Review with Meta-Analysis: Understanding Quality Characteristics of Edtech Interventions and Implementation for Disadvantaged Pupils* (Understanding Quality Characteristics of EdTech Interventions and Implementation for

Disadvantaged Pupils No. 2). Open Development & Education. https://doi.org/10.53832/opendeved.1115. Available from https://docs.opendeved.net/lib/G5P9ZNRI. Available under Creative Commons Attribution 4.0 International. (details)

Haßler, B., McBurnie, C., Walker, H., Klune, C., Huntington, B., & Bhutoria, A. (2024). *Protocol for a Systematic Review with Meta-Analysis: Understanding Quality Characteristics of Edtech Interventions and Implementation for Disadvantaged Pupils* (Understanding Quality Characteristics of EdTech Interventions and Implementation for Disadvantaged Pupils No. 1). Open Development & Education. https://doi.org/10.53832/opendeved.1077. Available from https://docs.opendeved.net/lib/2I2GT22T. (details)

He, M., Baker, R., Hutt, S., & Zhang, J. (2022). *A Less Overconservative Method for Reliability Estimation for Cohen's Kappa*. (details)

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied Statistics for the Behavioral Sciences* (5th ed). Houghton Mifflin ; [Hi Marketing] (distributor). http://catalog.hathitrust.org/api/volumes/oclc/50716608.html. (details)

Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, *2*(1), 119–135. https://doi.org/10.1007/s13194-011-0038-2.

Koçak, Ö., Göksu, I., & Göktas, Y. (2021). The factors affecting academic achievement: a systematic review of meta analyses. *International Online Journal of Education and Teaching*, *8*(1), 454–484. https://eric.ed.gov/?id=EJ1286675. Available from https://eric.ed.gov/?id=EJ1286675. (details)

Mayring, P. (2000). Qualitative content analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(2). https://doi.org/10.17169/fqs-1.2.1089. (details)

Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research*, *90*(1), 24–46. https://doi.org/10.3102/0034654319877153. (details)

Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Stansfield, C., Van Herwegen, J., Cottingham, S., & Higton, J. (2021). *What are the characteristics of teacher professional development that increase pupil achievement? Protocol for a systematic review* (pp. 1–37) [Report]. Education Endowment Foundation. https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/teacher-professional-development. (details)

Licences for this document

This document is available for reuse under either of the following licences.

Creative Commons Attribution 4.0

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Creative Commons Attribution 4.0 licence. To view this licence, visit https://creativecommons.org/licenses/by/4.0/deed.en.

The Open Government Licence v3.0

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0. To view this licence, visit https://nationalarchives.gov.uk/doc/open-government-licence/version/3 or email: psi@nationalarchives.gsi.gov.uk

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned. The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

This document is available for download at https://educationendowmentfoundation.org.uk.