A New Methodological Approach for Evaluating the Impact of Educational Intervention Implementation on Learning Outcomes

Laura A. Outhwaite[12*], Anthea Gulliford[1] & Nicola J. Pitchford[1]

[1]School of Psychology, University of Nottingham
[2]Centre for Education Policy and Equalising Opportunities, UCL Institute of Education

***Presenting author, l.outhwaite@ucl.ac.uk**
Other authors: anthea.gulliford@nottingham.ac.uk, Nicola.pitchford@nottingham.ac.uk

First choice conference section: Research Methods
Second choice conference section: Early Childhood Education

Background/Context:
Research shows that 1 in 5 randomised control trials (RCTs) in education include some form of implementation process evaluation (IPE; Connolly et al., 2018). However, there is limited connection between insights from the IPE to the impact evaluation which is vital for understanding how a specific intervention works in individual contexts. It can also provide vital insights into the scaling of interventions beyond feasibility/efficacy trials, where there are high levels of researcher control.

Purpose/Objective/Research Question:
The current IPE focuses on a maths app intervention that was empirically evaluated through a pupil level RCT (see Outhwaite et al., 2018). This IPE builds on the reported quantitative data analysis of learning outcomes and addresses two research questions:

RQ1: How was the maths app intervention implemented in individual participating schools?
RQ2: Is there a relationship between identified implementation themes and children's learning outcomes with the maths apps?

This pupil level RCT was conducted as a feasibility trial, prior to a larger scale efficacy trial of the same maths app intervention (see Nunes et al., 2019). As such, this study was exploratory. Previous research suggests that the introduction of technology alone into the classroom will not necessarily equal learning gains (Couse & Chen, 2010). Following, it was thought that individual school contexts and implementation narratives may be associated with how well children benefited from the intervention.

Setting:
The current study was conducted in 11 schools across the East Midlands, United Kingdom. The participating schools represented a range of socio-economic and multi-cultural backgrounds.

Population/Participants/Subjects:
A final sample of 389 children aged 4-5 years old took part in the pupil level RCT across 11 participating schools. Descriptive data for participating teachers was collected through a

self-report, end-of-project feedback questionnaire. The mean average number of years teaching experience for participating teachers was 9.92 (*SD* = 7.86). Teachers identifying as 'experienced' with technology represented 77% of the sample, while the remaining 23% identified as having 'limited experience'.

Intervention/Program/Practice:
The maths apps at the focus of this study have been developed by *onebillion*, an educational not-for-profit organisation. The apps are designed to support the acquisition of core basic mathematical concepts in Number, Shape, Space and Measure. Across the participating schools, children were randomly allocated to one of three groups: 1) the treatment group used the maths app intervention in addition to regular mathematics instruction, 2) the time-equivalent treatment group used the maths app intervention instead of a daily small group-based mathematics activity, thus time spent learning mathematics was equivalent to children in 3) the control group who continued to receive standard teacher-led mathematical instruction. In Group 1 and Group 2, the maths app intervention was implemented by classroom teachers for 30 minutes per day across 12 consecutive weeks. Children used the maths apps independently, with headphones, in a quiet area of their usual classroom environment.

Research Design:
The current study adopted a pragmatic, relativist, mixed-methods approach (Tashakkori & Teddlie, 2010) and was conducted in two phases relative to the two research questions. Phase 1 was primarily qualitative and employed an inductive, bottom-up approach to identify implementation themes in the data (Punch, 2013). Phase 2 drew upon a structured judgements methodology (Clarke, 2004) within a determinant theoretical framework (Nilsen, 2015).

To ensure the validity (authenticity) and reliability (trustworthiness) of the qualitative findings and consequent interpretations the following measures were taken: member checking, thick detailed descriptions, familiarity with participating schools, addressing demand effects, and audit trial.

Data Collection and Analysis:
Phase 1 drew upon data collected through narrative direct observations of intervention implementation across the 11 participating schools and self-report interviews with the participating teachers. Data was aggregated and coded using an inductive, bottom- up approach. Two cycles of open data coding were conducted to ensure the data was saturated and data codes were exclusive, exhaustive and consistent. A thematic analysis was then conducted to identify implementation themes.

Phase 2 utilised a structured judgements approach to generate a descriptive summary for each school that conveyed individual experiences of implementation the maths app intervention. Then a systematic quantitative framework (3-point Likert scale) was designed and applied to the descriptive summaries to assess the variability in implementation in each school. This procedure generated a quantitative index of intervention implementation for each school.

To examine the relationship between the implementation themes and children's learning outcomes in response to the maths app intervention, spearman's rho correlations were conducted between the quantitative index of intervention implementation and within-group effect sizes (Cohen's *d*) of children's learning gains. Any significant correlations were then inputted into a linear regression.

Findings/Results:
Results from Phase 1 identified four implementation themes; 'teacher support' (e.g. technical support and behavioural management), 'teacher supervision'(e.g. constant and consistent), 'intended implementation' (e.g. child had own iPad and calm classroom environment) and 'established routine' (e.g. consistent implementation timing and dedicated classroom space).

Results from Phase 2 showed a strong, positive and significant correlation between 'established routine' (M = 2.27, SD = .79) and effect sizes of children's learning gains (M = .77, SD = .36, r^s = .73, p = .011). No other significant correlations with learning gains were found. When entered into a linear regression analysis, 'established routine' significantly predicted children's learning outcomes with the maths apps (β = .29, p = .035), accounting for 41% of the observed variance, R^2 = .41, F (1,9) = 6.13, p = .035.

Conclusions:
This IPE study reports a novel and systematic methodological approach for understanding the relationship between variation in the implementation of a maths app intervention and its impact on associated learning outcomes. In this specific case, the importance of a well-established classroom routine was highlighted as an important factor when implementating the maths apps. This has significant implications for scaling this intervention within school settings to optimise effectiveness. However, it is important to acknowledge that the opportunity for variations in this specific intervention may be relatively minimal compared to other, more complex educational interventions. Nevertheless, this new methodological approach has the potential to be applied to other educational interventions and should be trialled in other contexts.

References

Clarke, D. 2004. "Structured Judgement Methods – the Best of Both Worlds?" In *Mixing Methods in Psychology: The Integration of Qualitative and Quantitative Practice*, edited by Z. Todd, B. Nerlich, S. McKeown, and D. Clarke, 79–99. Hove: Psychology Press.

Connolly, P., C. Keenan, and K. Urbanska. 2018. "The Trials of Evidence-based Practice in Education: A Systematic Review of Randomised Controlled Trials in Education Research 1980–2016." *Educational Research* 60 (3): 276–291.

Couse, L. J., and D. W. Chen. 2010. "A Tablet Computer for Young Children? Exploring its Viability for Early Childhood Education." *Journal of Research on Technology in Education* 43 (1): 75–96.

Nilsen, P. 2015. “Making Sense of Implementation Theories, Models and Frameworks.” *Implementation Science* 10 (53): 1–13.

Nunes, T., L. Malmberg, D. Evans, D. Sanders-Ellis, S. Baker, R. Barros, P. Bryant, and M. Evangelou. 2019. *onebillion Evaluation Report.* London: Education Endowment Foundation.

Outhwaite, L. A., M. Faulder, A. Gulliford, and N. J. Pitchford. 2018. “Raising Early Achievement in Math with Interactive Apps: A Randomized Control Trial.” *Journal of Educational Psychology* 111 (2): 284–298.

Punch, K. F. 2013. *Introduction to Social Research: Quantitative and Qualitative Approaches*. London: Sage.

Tashakkori, A., and C. Teddlie. 2010. “Putting the Human Back in ‘Human Research Methodology’: The Researcher in Mixed Methods Research.” *Journal of Mixed Methods Research* 4 (4): 271–277.