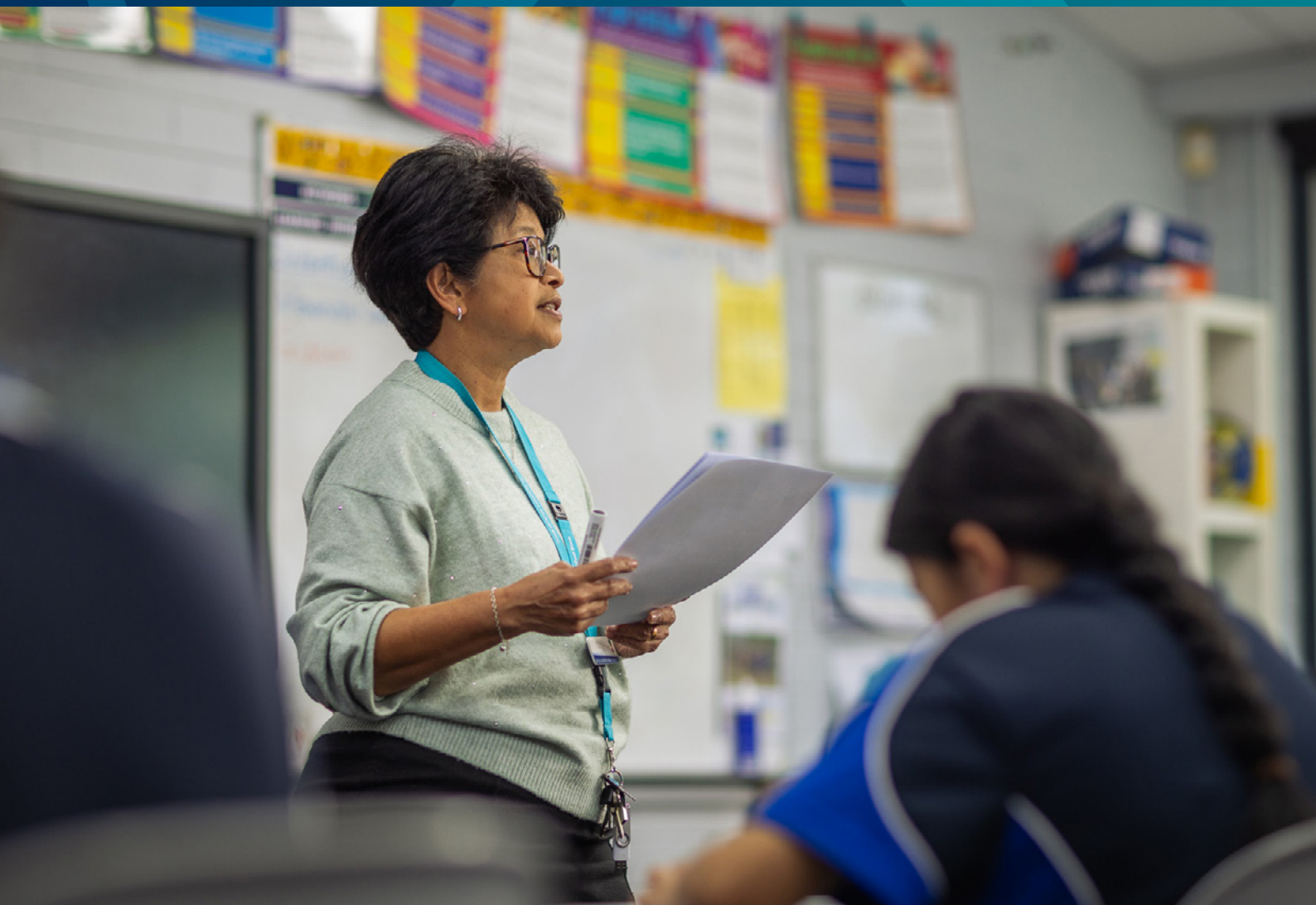
Australian Education Research Organisation

The impact of context on evidence-based practices

A rapid literature scan on formative assessment, explicit instruction and mastery learning

September 2024



The Australian Education Research Organisation (AERO) is Australia's national education evidence body, working to achieve excellence and equity in educational outcomes for all children and young people.

Acknowledgement

AERO's work is made possible by the joint funding it receives from Commonwealth, state and territory governments.

AERO would like to acknowledge the Expert Review Panel of the Evidence-Based Practices in Schools project for guidance provided in relation to this research. We especially thank Dr Nathaniel Swain, whose feedback supported improved framing, structure, terminology of the research report and enhanced description of our review in relation to a broader program of work.

Acknowledgement of Country

AERO acknowledges the Traditional Custodians of the lands, waterways, skies, islands and sea Country across Australia. We pay our deepest respects to First Nations cultures and Elders past and present. We endeavour to continually value and learn from First Nations knowledges and educational practices.

Authors

Zid Mancenido, Karen Gonsalkorale, Nuella Flynn and Jessica Brosnan

How to cite

Mancenido, Z., Gonsalkorale, K., Flynn, N., & Brosnan, J. (2024). *The impact of context on evidence-based practices: A rapid literature scan on formative assessment, explicit instruction and mastery learning*. Australian Education Research Organisation. https://www.edresearch.edu.au/research/research-reports/impact-context-evidence-based-practices-rapid-literature-scan

Publication details

ISBN 978-1-923066-43-4 (online)
Cover image: AERO

Contents

Introduction

Evidence-based practices are backed up by research evidence. This means there is broad consensus from rigorously conducted evaluations that they work. The Australian Education Research Organisation (AERO)'s Standards of Evidence establish our view on what constitutes rigorous and relevant evidence. When evidence is determined to be rigorous and relevant, it provides confidence that a particular approach will be effective in a particular context and better support decision-making that is evidence-informed.

We conducted a literature review of whether studies of effectiveness of 3 well-established, evidence-based teaching practices reveal variation in positive impact across contexts. Specifically, we examined meta-analytic reviews of the effects of formative assessment, explicit instruction and mastery learning on student achievement, and whether these studies showed that effects vary depending on contextual factors such as location, subject area, year level and student demographics.

When considering education evidence, teachers, educators and policymakers may wonder how applicable it is to the context they work in. This may prompt questioning about whether research conducted in a dissimilar context – for example, a different location, student demographic, year level or subject area – can offer insights that are more broadly relevant.

Our goal was to better clarify if context matters for the effectiveness of these practices. This report presents the methodology and findings of our investigation.

How we use 'context'

Throughout this report, we refer to context, contextual variables and contextual factors interchangeably when referring to the **location, subject area, year level and student demographics** of different studies that were reviewed in the meta-analyses.

We examined the evidence base for 3 evidence-based teaching practices – formative assessment, explicit instruction and mastery learning – to investigate whether the studies showed variation in effectiveness depending on contextual factors such as **location, subject area, year level and student demographics**.

We aimed to better specify whether studies that examined the effectiveness of these practices revealed differences related to context, and to support teachers, educators and policymakers in better identifying whether evidence for teaching practices is relevant to their context.

Key findings

- Formative assessment, explicit instruction and mastery learning are effective approaches that consistently have a positive impact on student achievement. The meta-analyses we reviewed combined and synthesised studies conducted over a wide range of contexts, representing various locations, subject areas, year levels and student demographics.
- There is little evidence in these studies that effect sizes of formative assessment, explicit instruction and mastery learning vary across different contexts. This suggests that formative assessment, explicit instruction and mastery learning are beneficial across different subgroups of students and subject areas.
- Although we found that formative assessment, explicit instruction and mastery learning have a positive impact on student achievement across different contexts, many of the meta-analyses reviewed did not directly analyse the impact of context or examined relatively few contextual variables. Further research focused specifically on the effect of context, as well as on Australian contexts and student demographic variables within Australia could nuance our understanding of whether and how context influences the effectiveness of teaching practices.

Insights on practices

Overall, findings indicate that formative assessment, explicit instruction and mastery learning are beneficial across different subgroups of students and subject areas.

Studies conducted across various locations provided the following insights:

Formative assessment:

- has a positive impact on student achievement in mathematics, reading, writing, social science and languages
- works for primary and secondary students
- benefits students with and without additional learning needs.

Explicit instruction:

- has a positive impact on student achievement in mathematics, reading, spelling, problem-solving and science
- works for primary and secondary students
- benefits students with and without additional learning needs.

Mastery learning:

- has a positive impact on achievement in mathematics, sciences, social studies and English and languages
- works for primary and secondary students
- is particularly effective for lower-achieving students, with higher-achieving students also benefitting.

Methodology

We examined meta-analytic reviews of the impact of formative assessment, explicit instruction and mastery learning on student achievement. Each review analysed a set of studies on one of the 3 teaching practices to calculate an overall effect size for that practice. Some reviews additionally reported (whether quantitatively or through narrative review) whether effect sizes varied as a function of contextual variables.

To provide an overview of the effectiveness of the 3 teaching practices, we report findings relating to overall effect sizes. We summarise the contexts of the studies reviewed in the meta-analyses by describing study location, subject area and sample characteristics (for example, student demographics, year level or learning stage). Where possible, we also report whether effect sizes varied as a function of contextual factors and the relationship between contextual variables and effect sizes.

In education research, effect sizes quantify the strength or magnitude of the effect of a particular intervention (such as a policy, program or practice) on an outcome of interest (such as student achievement or learning over time). Standardised mean differences are a commonly used effect size, and they express the difference in outcomes between the intervention group and comparison group in standard deviation units. This standardisation allows for comparability of like effect sizes across studies. In a meta-analysis, effect sizes from the included studies are combined using a statistical model to calculate an overall effect size (Cooper & Hedges, 2009). Throughout the report, we classify effect sizes less than -0.20 as negative, effect sizes between -0.19 and 0.20 as null effects, and greater than 0.20 as positive (Cohen, 1988).

We have selected meta-analyses that include rigorously designed studies. Typical inclusion criteria stipulated those studies had to be experiments or quasi-experiments, contain a control group and be free from major methodological flaws. In other words, the studies were designed to test whether the teaching practice caused positive effects (at least Level 3 according to AERO's Standards of Evidence).

Studies in the meta-analyses were conducted in classroom settings. Many of the meta-analyses included both school-aged and tertiary-aged samples. Wherever possible, we restricted our analysis to effect sizes for school-aged students to enhance the relevance of our findings.

In most studies, student achievement was measured as performance on locally developed exams assessing the material or skills taught.[1] Examples include performance on mathematics or reading exams.

To identify relevant meta-analyses, we searched the Education Endowment Foundation's (EEF) Teaching and Learning Toolkit (EEF, 2021) and the Visible Learning MetaX Global Research Database. We supplemented these sources with literature searches to identify additional meta-analyses.

1 Effect sizes on student achievement tend to be smaller when system-wide standardised tests are used as outcome measures as opposed to teacher- or school-developed exams (Slavin, 1990). This is to be expected, as teacher- or school-developed exams are generally more closely aligned to what students are taught through the evidence-based practice.

Limitations

Thousands of empirical studies have been conducted on explicit instruction, mastery learning and formative assessment. We knew that it would be beyond our resources to review all of them to assess the potential influence of context, so we conducted an umbrella review of the numerous meta-analyses that have already been undertaken on the large corpus of research on these practices. Further, we chose to narrow our focus to meta-analyses because they provide a quantitative estimate of the extent to which practices work, making it easy to compare findings across reviews.

There are constraints and trade-offs that must be made in any methodology, which in turn limit the claims that can be made. We did not conduct an exhaustive literature review, nor review individual studies.

These methodological choices mean that qualitative studies and narrative reviews were not included in our review. This is a significant limitation of our report, particularly, for example, as more recent educational research into culturally responsive teaching has primarily used qualitative or critical methodologies (Bottiani et al., 2018). AERO's research into cultural responsiveness notes that there is currently limited causal evidence regarding the effectiveness of culturally responsive practices on student outcomes. Instead, further research and clearer evidence-based advice about implementing culturally responsive practices and behaviours integrated with explicit teaching approaches holds the promise of accelerating improved outcomes for First Nations students, as well as benefiting all students.

Our methodological choices also mean it is possible we did not identify or review all relevant meta-analyses. Given that all meta-analyses reviewed seemed to hold minimal variability in their findings, we did not think that accidental exclusion was likely to dramatically influence our conclusions.

Finally, reviewing only the meta-analyses and not individual studies meant we were reliant on authors' reports on contextual variables. Sometimes, these contextual variables could not be easily extracted, and other times, as is often the case when sample sizes are large, there was variability in study designs, specificity of the interventions (particularly for formative assessment and explicit instruction), outcomes measures and effect sizes. Thus, it was sometimes difficult to pinpoint the precise conditions under which the evidence-based practices were deemed effective, and for whom, and researcher judgments were made in order to harmonise and compare contextual influences across meta-analyses. To minimise the subjectivity of these judgments, all classification decisions were made by a pair of researchers with different research backgrounds, and consensus decisions were reached before proceeding.

Analysing how context influences the effectiveness of key teaching practices

This section summarises the influence of context on the effectiveness of 3 evidence-based teaching practices: formative assessment, explicit instruction and mastery learning.

Formative assessment

We identified 6 meta-analyses reviewing 138 unique studies. The reviews, published between 1986 and 2021, covered studies from 1974 to 2020.

Studies examined the effectiveness of gathering assessment information that allows teachers to provide feedback and adjust their instruction to meet student learning needs. In some cases, feedback was delivered by peers or computers instead of teachers. Consistent with Black and Wiliam's (1998) conceptualisation of formative assessment, these practices are described in AERO's Formative Assessment practice guide.

Although some studies included all 3 key features of formative assessment (gathering assessment information, giving feedback and adjusting instruction), others focused on specific features. For example, some studies examined the effectiveness of feedback within a formative assessment intervention but did not report whether teachers adjusted their instruction in response to student learning data.[2] Where possible, we report findings related to specific features of formative assessment rather than to formative assessment as a whole.

Overall findings

The 6 meta-analyses reported overall effect sizes ranging from 0.20 to 0.72. Formative assessment practices resulted in significantly greater student achievement compared to control groups, and overall they were effective regardless of contextual factors.

Regarding the key features of formative assessment, one meta-analysis (Lee et al., 2020) examined the relationship between types of instructional adjustments and student achievement. Planning ahead for how to adjust instruction in response to assessment data and then making those adjustments predicted higher student achievement. However, effect sizes were not significant for formative assessment interventions involving unplanned adjustments or no adjustments. These findings suggest that planned instructional adjustments are an important feature of effective formative assessment interventions.

A meta-analysis focusing on writing (Graham et al., 2015) reported an effect size of 0.61 for studies focusing on feedback as a feature of formative assessment. Studies focusing on progress monitoring produced an effect size of 0.18.

2 Where possible, studies that focused just on giving feedback without the initial gathering of assessment information were excluded as they do not fully meet the conceptualisation of formative assessment as described in AERO's practice guides.

Study location

Three meta-analyses (Fuchs & Fuchs, 1986; Graham et al., 2015; Kingston & Nash, 2011) did not report study location but only reviewed studies published in English. Two meta-analyses (Klute et al., 2017; Lee et al., 2020) restricted all or most studies to those conducted in the United States. Karaman (2021) included studies conducted in Turkey only.

The overall effect size (0.72) for studies in Turkey (Karaman, 2021) was larger than the effect sizes for the predominantly United States samples (0.26 from Klute et al., 2017; 0.29 from Lee et al., 2020). However, the studies in the Turkish meta-analysis were all conducted after educational reform placed a strong emphasis on formative assessment practices. It is unclear to what extent the findings reflect cross-national differences, factors linked to educational reform or other unknown factors.

The meta-analyses did not report whether the studies were conducted in inner-city, regional or rural areas.

Subject area

Overall, findings indicate a beneficial effect of formative assessment for mathematics, literacy and arts. More research is needed to understand the effect of formative assessment on science achievement.

Most reviews covered a variety of subject areas, including mathematics, reading, writing, science, social science and languages. One meta-analysis focused on studies of writing.

Three meta-analyses reported effect sizes as a function of subject area. Lee et al. (2020) found effect sizes of 0.34 for mathematics, 0.33 for literacy and 0.29 for arts. Klute et al. (2017) reported a larger effect size for mathematics (0.36) than for reading (0.22) and writing (0.21). In contrast, Kingston and Nash (2011) found a larger effect size for English language arts (0.32) than for mathematics (0.17).

Effect sizes for science (0.13 and 0.09) have been smaller than for other subject areas (Kingston & Nash, 2011; Lee et al., 2020). However, this result should be interpreted with caution, as there were fewer studies focusing on science and, therefore, included across the meta-analyses (for example, Klute et al., 2017 noted they were unable to include any science studies in their meta-analysis as these studies did not meet the US Department of Education's What Works Clearinghouse standards).

Learning stage

Findings indicate that formative assessment benefits students across all educational stages.

Four of the 6 meta-analyses restricted the samples to school-aged students. One meta-analysis included all school-aged samples except for one preschool sample. Only one meta-analysis reported findings from both school-aged and tertiary samples.

Four meta-analyses (Fuchs & Fuchs, 1986; Kingston & Nash, 2011; Lee et al., 2020; Karaman, 2021) calculated effect sizes for subgroups of students as a function of learning stage. Whether the authors compared primary to secondary samples (for example, Lee et al., 2020) or conducted a more granular analysis of year levels (for example, Kingston & Nash, 2011), none reported significant moderation by educational level. Effect sizes were positive for all subgroups.

Student demographics

None of the meta-analyses focused on student demographic variables such as gender, socio-economic status or ethnic background.

Students with additional learning needs

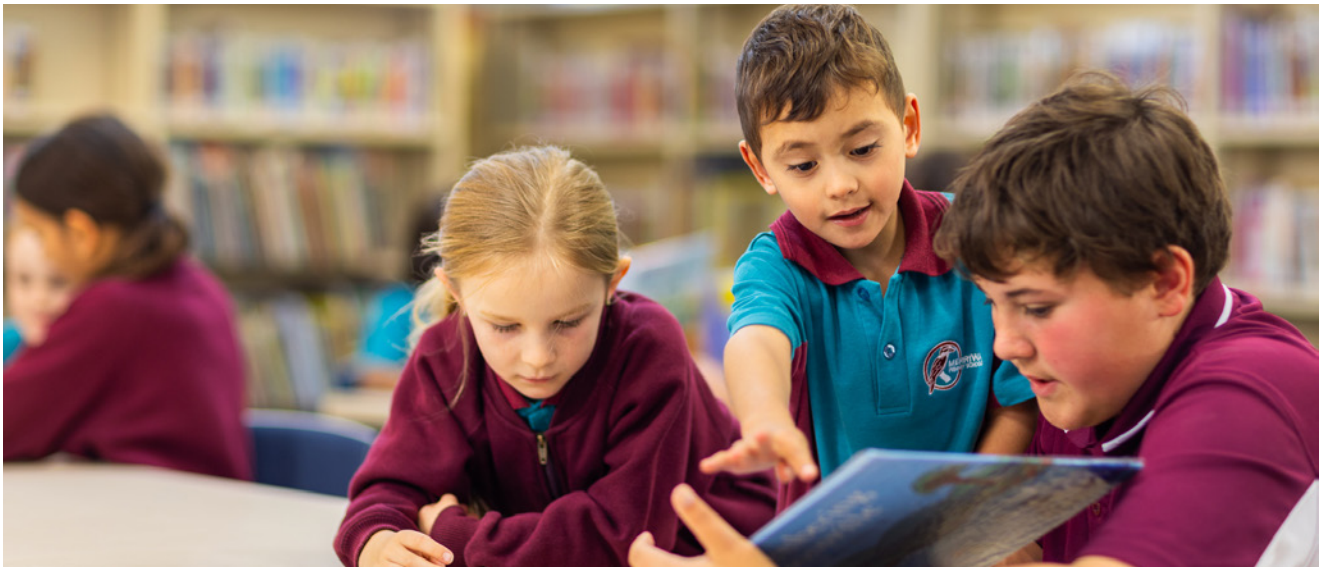
Findings suggest that formative assessment is effective regardless of students' additional learning needs. The effects of formative assessment were not significantly different for students diagnosed with an intellectual disability than for students without an intellectual disability (Fuchs & Fuchs, 1986), or for mainstream classrooms versus special education classrooms (Lee et al., 2020). The effect sizes were positive and significant for the different subgroups of students.

Summary for formative assessment

Overall, findings indicate that formative assessment has a significant, positive effect on student achievement. The studies spanned the 1970s to the 2020s, with many studies conducted in the last 2 decades. Although many of the formative assessment studies reviewed in the meta-analyses were conducted in English-speaking countries, recent studies from Turkey were also represented. The Turkish studies yielded a higher effect size, on average, than meta-analyses of predominately United States studies. However, it is not clear how to best interpret this difference.

Formative assessment is beneficial for mathematics, literacy and arts, although it is not clear whether it is effective for science. There is little evidence that the effectiveness of formative assessment depends on educational level or the additional learning needs of students. However, few conclusions can be drawn for other student demographic variables.

In sum, the meta-analyses examined relatively few contextual variables. However, when they were examined, there was little evidence that the effectiveness of formative assessment varied across contextual factors. Overall, these findings suggest that formative assessment is beneficial for a variety of students across different contexts.



Explicit instruction

We reviewed 12 meta-analyses, compiling more than 328 unique studies in total.[3] The reviews were published between 1990 and 2018, with the studies stretching from 1972 to 2017.

Studies in the meta-analysis drew on models of explicit instruction (for example, Archer & Hughes, 2011; Klahr, 2009) that emphasise clear teacher guidance through the learning process. Typical explicit instruction interventions involve teachers stating the purpose and rationale for learning the new skill, clearly explaining and demonstrating what students need to learn and providing opportunities for practice. Although the study interventions did not necessarily include all of these features of explicit instruction, they were broadly consistent with the practices described in AERO's Explicit Instruction practice guide.

Direct Instruction refers to scripted programs based on the work of Engelmann (1967; 2007) that emphasise curriculum design rather than teacher behaviours. We included one meta-analysis of Direct Instruction that reported findings on context based on a large sample size because it is also a fully guided instructional approach.

We did not analyse meta-analyses involving predominantly lab-based interventions or tertiary-aged students. This led to the exclusion of meta-analyses focusing on second language acquisition.

Overall findings

Effect sizes reported in the meta-analyses ranged from 0.40 to 1.22, indicating that explicit instruction has a positive impact on student achievement.

As outlined next, there was little evidence overall to suggest that the effectiveness of explicit instruction depends on context.

Study location

The meta-analyses focused on studies published in English. Most did not specify where the studies were undertaken or whether effect sizes varied based on location. Information was more readily available for locations in the United States than for other locations (for example, Swanson & Hoskyn, 1998).

In their analysis of 328 Direct Instruction studies, Stockard et al. (2018) noted the studies were conducted in urban, regional and rural areas in the United States, as well as in other countries. Meta-regression results indicated 2 significant results for location. Specifically, effect sizes for urban United States locations were higher for ability measures but smaller for language achievement than for other locations. The other locations included suburban United States, rural United States and other countries (with type of area not specified). It is difficult to interpret this finding.

3 This is a conservative estimate. The possible range of unique studies is between 328 (the largest number of studies reviewed in a single meta-analysis) and 765 (the sum of the studies reviewed across all meta-analyses). It was difficult to identify how many of the 765 studies were reviewed in more than one meta-analysis.

Subject area

Findings suggest that explicit instruction is an effective instructional method for a range of subject areas, with some variation in the magnitude of the effect by subject area.

Five meta-analyses focused on mathematics, 2 on writing, one on critical thinking and one on the integration of texts. Three covered studies on a range of subject areas. Irrespective of their subject area focus, all meta-analyses found significant and positive effects of explicit instruction overall.

Further analyses of effect sizes by subject area showed there was some variation in the degree of the effect reported for explicit instruction. Methodological differences may partly explain the variation in effectiveness across subject area. Stockard et al. (2018) tested the impact of Direct Instruction, whereas Alfieri et al. (2011) used a broader definition of explicit instruction and compared its effects to unassisted discovery-based learning. Alfieri et al. (2011) found effect sizes for explicit instruction of 0.95 for verbal and social skills, 0.48 for problem-solving skills, 0.39 for science and 0.16 for mathematics. This would appear to suggest that explicit instruction is most effective for verbal and social learning tasks, followed by problem-solving, science and mathematics. In contrast, subject area did not significantly predict the magnitude of effect sizes for Direct Instruction in Stockard et al. (2018). They reported effect sizes of 0.55 for maths, 0.51 for reading, 0.66 for spelling, 0.54 for language and 0.41 for studies of other subject areas or multiple subject areas. Given Alfieri et al. (2011) found greater variation across subject areas, we interrogated the 5 meta-analyses focused on mathematics to understand more about the effectiveness of explicit instruction. Across the 5 studies, we found effect sizes that ranged from 0.55 to 1.22, indicating that explicit instruction has a strong positive impact on student achievement in mathematics.

Year level

Findings suggest that explicit instruction benefits students irrespective of their year level. Even when age or year level differences in degrees of effectiveness have been found, effect sizes for all subgroups were significant and positive.

Half of the meta-analyses focused exclusively on school-aged students and the other half comprised samples of predominantly school-aged students. Alfieri et al. (2011) examined effect sizes for children (less than 12 years old or up to Year 6), adolescents (aged 12 to 17 or Years 7 to 12) and adults (greater than 18). Effect sizes for explicit instruction were significant and positive for all 3 age groups. A post-hoc comparison found that adolescents (effect size of 0.53) benefitted more from explicit instruction compared to unassisted discovery learning than adults (effect size of 0.26).[4]

Two meta-analyses found no evidence that year level moderated results. One of these analyses (Jacobse & Harskamp, 2011) found no significant difference between students in Kindergarten to Year 3, and Years 4 to 6. In their analysis of mostly pre-Year 4 samples, Stockard et al. (2018) found that year level did not significantly predict effect sizes of Direct Instruction.

4 The authors did not report the effect size for children or comparisons with the effect size for adolescents or adults.

Student demographics

Little demographic data was available for our review. Based on the few meta-analyses that reported student demographic variables, the samples included male and female students from both ethnic minority (for example, African American) and majority (European American) groups. One meta-analysis (Gersten et al., 2009) reported that the majority of students were from a lower socio-economic background, and another (Stockard et al., 2018) reported that a quarter of the studies focused on students from high-poverty backgrounds.

Neither socio-economic status nor ethnic background consistently predicted effect sizes of Direct Instruction (Stockard et al., 2018). Direct instruction was effective irrespective of these demographic variables.

Students with additional learning needs

Reviews found significant positive effect sizes of explicit instruction for students both with and without additional learning needs.

Half of the meta-analyses focused exclusively on 'low-achieving' students or students identified as having a 'learning difficulty'. For the purpose of this summary, we refer to these groups collectively as 'students with additional learning needs'. These meta-analyses reported overall effect sizes ranging from 0.58 to 1.22.

Two of the meta-analyses included samples of students with and without additional learning needs. Jacobse and Harskamp (2011) found that explicit instruction had stronger effects in special education classrooms than in mainstream classrooms. On the other hand, Stockard et al. (2018) reported that 'at-risk' status did not significantly predict the effectiveness of Direct Instruction. Students were classified as 'at-risk' if they received special education services or were described as 'remedial' or 'low-achieving'.

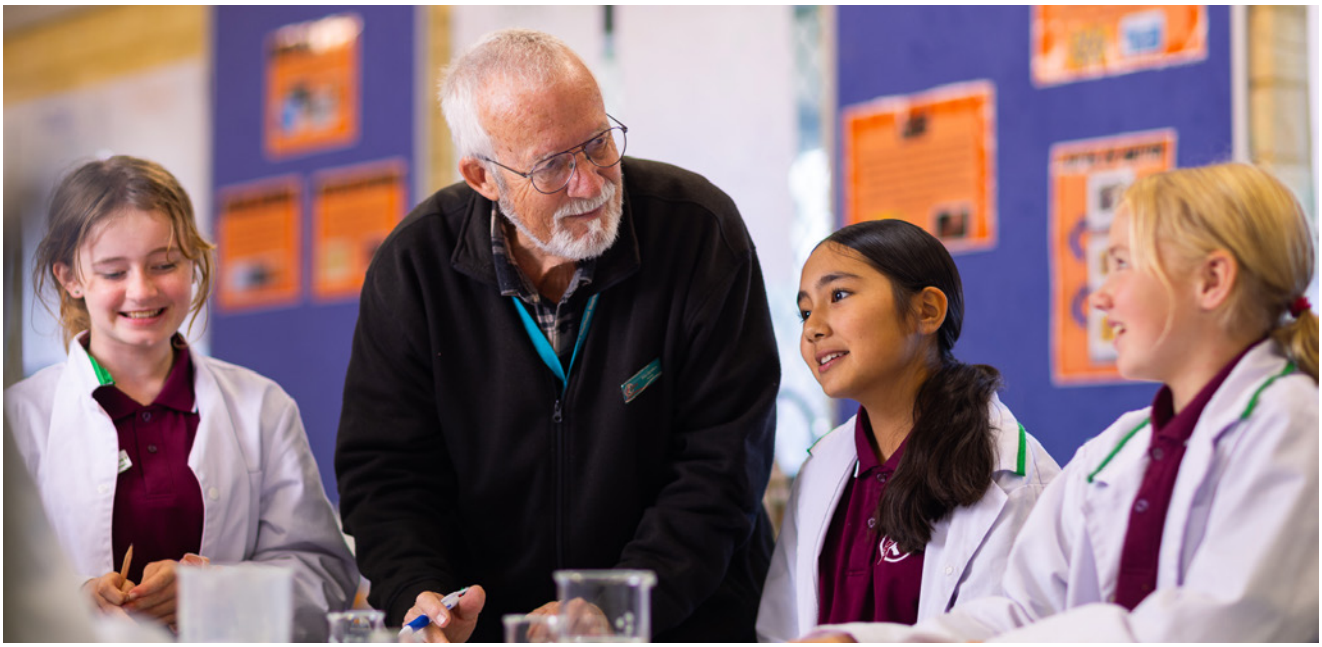
It is important to note that authors classified students according to different criteria (for example, based on test or ability scores, whether students had an Individualised Education Plan or received special education). Students classified as having additional learning needs were a heterogeneous group.

Although specific interventions may be required according to the type of additional learning need, the positive effect sizes suggest that students with and without additional learning needs benefit from explicit instruction.

Summary for explicit instruction

The meta-analyses demonstrated the efficacy of explicit instruction on student achievement across a range of contexts. There was no consistent evidence that context moderated the positive impact of explicit instruction. Findings suggest that explicit instruction is beneficial across subject areas and year levels, and for students with and without additional learning needs.

It should be noted that few meta-analyses reported on contextual variables, and at times, there was inconsistency in how these variables were coded. Future research addressing these gaps would enhance understanding of the impact of explicit instruction across different contexts.



Mastery learning

We identified 10 meta-analytic reviews for our analysis. The reviews were published between 1976 and 2020, with the majority in the 1970s and 1980s. The mastery learning meta-analyses focused on studies of school-aged children.

In most studies, students in the mastery learning group were required to achieve a high level of proficiency before moving on to new content. Some studies examined group-based, teacher-paced mastery learning interventions. This operationalisation fits with Bloom's (1968) Learning for Mastery. In other studies, mastery learning was self-paced, in line with Keller's (1968) Personalised System of Instruction. A meta-analysis (Kulik et al., 1990) found no significant difference in the effectiveness of group-based versus self-paced approaches to mastery learning. For our analysis, we focused on effect sizes for group-based, teacher-paced mastery learning as this is consistent with the approach to mastery learning as described in AERO's Mastery Learning practice guide.

The meta-analyses also presented study data in a way that enabled us to extract study-level findings. Thus, we report findings at the study level, rather than at the meta-analysis level (as we did for formative assessment and explicit instruction).

Overall findings

A total of 87 effect sizes (from 81 studies[5]) were included in our review. The overwhelming majority (78%) of these effect sizes were positive. Mastery learning interventions resulted in significantly greater student achievement compared to control groups.

As outlined next, in general, overall mastery learning was effective regardless of contextual factors.

5 The number of effect sizes is larger than the number of studies because some studies contained multiple samples and/or reported effect sizes for more than one measure of student achievement.

Study location

Although many studies sampled students in North America, other studies sampled students in England, Chile, Israel or Nigeria. We did not find any studies conducted in Australia. There was representation from inner-city, regional and rural schools across the studies.

We were unable to extract location data for all samples. However, based on the available data, effect sizes did not appear to vary as a function of location.

Subject area

Findings suggest that overall, mastery learning is effective across a range of subject areas.

Subjects examined in mastery learning studies included mathematics, biology, chemistry, physics, reading, languages, anthropology, history and questioning skills. In line with Guskey and Gates (1986), we grouped subjects into the following categories: mathematics, sciences, language arts (English and languages) and social studies (government, history, general social studies).

The 85 effect sizes available for the subject area analysis were distributed as follows: mathematics (54%), sciences (14%), language arts (20%) and social studies (12%). The percentage of positive effect sizes was similar across the different subject areas, ranging from 70% to 80%.

Educational stage

Findings suggest that mastery learning is effective regardless of educational stage.

We extracted educational stage data for 84 reported effect sizes. Of these effect sizes, 39% were derived from primary school samples and 61% were from secondary school samples. The effect sizes were positive for 85% of primary school samples and 76% of secondary school samples.

Student demographics

Overall, there is limited research on the effectiveness of mastery learning as a function of gender, socio-economic or ethnic background.

Samples included male and female students from different socio-economic groups, as well as ethnic minority and majority groups. These student demographic variables have not been the focus of the meta-analytic reviews. We found only 3 studies published in the 1970s and one in 1993 reporting on findings as a function of these student demographic variables.

One study (Dolan, 1993, as cited in EEF, 2021) reported that mastery learning had a significant effect on male, but not female, students. In another study, mastery learning was more effective for high-achieving girls than for low-achieving girls. The reverse pattern was found for boys, with low-achieving boys benefitting more from mastery learning than high-achieving boys.

One study (Saunders-Harris & Yeany, 1981, as cited in EEF, 2021) found no evidence that the effect of mastery learning varied according to ethnic group status.

Findings were inconclusive for the one study (Kersch, 1970, as cited in EEF, 2021) examining whether the effect of mastery learning varied according to students' socio-economic status.

Student aptitude and prior achievement scores

Overall, studies indicate that mastery learning is effective for students regardless of their aptitude, but it has a greater impact on students lower in aptitude than students higher in aptitude.

Nine studies reported whether student aptitude or prior achievement scores moderate the impact of mastery learning on student achievement. Six of these studies found that mastery learning was more effective for students with lower aptitude than for students with higher aptitude (Burrows & Okey, 1975; Dillashaw & Okey, 1983; Hymel & Gaines, 1977; Mevarech, 1986, 1991 and Wyckoff, 1974, as cited in Kulik & Kulik, 1986; Kulik et al., 1990 and EEF, 2021). One study (Jones, 1975, as cited in Kulik & Kulik, 1986) found the reverse effect, and 2 studies (Arblaster, 1991 and Mevarech, 1986, as cited in EEF, 2021) found no significant moderation by student aptitude.

Other studies examined the correlation between student aptitude and achievement in mastery and non-mastery conditions. Two studies (Anania, 1981 and Nordin, 1970, as cited in EEF, 2021) found that the correlation between aptitude and achievement was weaker in the mastery condition than in the non-mastery condition. Two studies (Dunkelberger & Heikkinen, 1984 and Fagan, 1976, as cited in Kulik & Kulik, 1986 and in EEF, 2021) found moderate correlations between student aptitude and achievement in both conditions.

Importantly, most studies (10 out of the 13) found that mastery learning had a positive effect on student achievement regardless of student aptitude.

Summary for mastery learning

Mastery learning studies have been conducted with students from a range of backgrounds in various locations. There was a high representation of studies conducted in North America and with maths as the subject area. By and large, findings indicate that mastery learning has a significant positive effect on student achievement.

There is little evidence that the effectiveness of mastery learning among school-aged samples depends on educational stage and subject area. Although students lower in aptitude appear to benefit more from mastery learning, students higher in aptitude also benefit.

Further research is needed in order to form conclusions about the effectiveness of mastery learning on specific student demographic variables such as socio-economic status.

Future research could expand the evidence base to include more studies conducted outside of the United States and focus on student demographic variables such as socio-economic status. This would also update the evidence base, which consists of many studies conducted before the 1990s.

Notwithstanding these caveats, our review of the evidence base suggests that mastery learning does indeed work, for a variety of students across different contexts.

Conclusion

Multiple meta-analytic reviews of numerous studies have found that formative assessment, explicit instruction and mastery learning are effective teaching practices. These 3 evidence-based practices align with our understanding of the learning process common to students across diverse learning contexts.

Studies of these evidence-based practices have been conducted over a wide range of contexts, representing various locations, student demographics, year levels/learning stages and subject areas.

There was little evidence that effect sizes for the 3 evidence-based practices varied across contexts. Overall, findings indicate that formative assessment, explicit instruction and mastery learning are beneficial across different subgroups of students and subject areas.

Studies conducted across various locations provided the following insights:

Formative assessment:

- has a positive impact on student achievement in mathematics, reading, writing, social science and languages
- works for primary and secondary students
- benefits students with and without additional learning needs.

Explicit instruction:

- has a positive impact on student achievement in mathematics, reading, spelling, problem-solving and science
- works for primary and secondary students
- benefits students with and without additional learning needs.

Mastery learning:

- has a positive impact on achievement in mathematics, sciences, social studies and English and languages
- works for primary and secondary students
- is particularly effective for lower-achieving students, with higher-achieving students also benefitting.

These findings show that formative assessment, mastery learning and explicit instruction are likely to work in a variety of classrooms.

Research focused specifically on context could help us further understand whether and how context influences the effectiveness of teaching practices. Many of the meta-analyses did not directly analyse the impact of context, while others examined relatively few contextual variables.

More recently, AERO has expanded on the research findings in this report through our model of learning and teaching that identifies the most effective and efficient teaching practices, aligned with how students learn at various phases of the learning process. The model recognises that all students benefit from evidence-based practices that align with the processes of acquiring, retaining, retrieving and consolidating learning. Studies into the effectiveness of formative assessment, explicit instruction and mastery learning in an Australian context are limited. Our model recognises that other conditions influence the effectiveness of these approaches and includes additional practices to plan for and enable a safe, supportive and learning-focused environment to meet the learning needs and aspirations of all students.

It is unlikely that research findings will ever be generated in contexts that are precisely the same as a teacher or educator's own context. This means that they will often face the challenge of how best to identify whether evidence for teaching practices is relevant to their context.

Further reading on using research evidence

AERO has developed evidence decision-making tools based on our Standards of Evidence to help teachers, educators and policymakers assess the relevance of research evidence and make informed decisions about new or existing policies, programs or other initiatives. Given the limited research in Australian contexts and on Australian student cohorts, these tools may be especially beneficial in considering when and how to apply international research in Australian schools. These tools help schools to:

- assess confidence as to whether a certain practice or program is likely to be effective in that context
- decide on next steps, including how to implement the practice or program given the level of confidence
- collect additional evidence to further consider and monitor effectiveness.

AERO has also developed a model of learning and teaching underpinned by the approaches explored in this report, bringing together research on the student learning process with practical implications for teachers. This model helps to explore evidence-based teaching practices through their alignment with key aspects of the learning process. AERO's model also recognises the importance of teachers developing responsiveness to create learning environments that are culturally safe and support responsive practices to meet the learning needs and aspirations of all students. The model can support teachers in making decisions and applying particular techniques and strategies in their context with greater understanding of how they may optimise student learning.

Once schools have enough confidence in the evidence to embed evidence-based approaches in their teaching, AERO's series of explainers provides advice for taking a deliberate and structured approach to implementation. These explainers highlight the importance of considering context at the earliest stages of implementing evidence-based practices. This helps schools make implementation decisions based on what is most likely to work to improve outcomes for students in their context.

References

Bottiani, J. H., Larson, K. E., Debnam, K. J., Bischoff, C. M., & Bradshaw, C. P. (2018). Promoting educators' use of culturally responsive practices: A systematic review of inservice interventions. *Journal of Teacher Education*, *69*(4), 367–385. https://doi.org/10.1177/0022487117722553

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.

Cooper, H., & Hedges, L. V. (2009). Research synthesis as a scientific process. In H. Cooper, L. V. Hedges & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 3–16). Russell Sage Foundation. http://www.jstor.org/stable/10.7758/9781610441384.5

Formative assessment

Meta-analyses included in review

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*(3), 199–208. https://doi.org/10.1177/001440298605300301

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *Elementary School Journal*, *115*(4), 523–547. https://doi.org/10.1086/681947

Karaman, P. (2021). The effect of formative assessment practices on student learning: A meta-analysis study. *International Journal of Assessment Tools in Education*, *8*(1), 801–817. https://doi.org/10.21449/ijate.870300

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, *30*(4), 28–37. https://doi.org/10.1111/j.1745-3992.2011.00220.x

Klute, M., Apthorp, H., Harlacher, J., & Reale, M. (2017). *Formative assessment and elementary school student academic achievement: A review of the evidence* (REL 2017–259). Regional Educational Laboratory Central.

Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K–12 education: A systematic review. *Applied Measurement in Education*, *33*(2), 124–140. https://doi.org/10.1080/08957347.2020.1732383

Other references cited

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Explicit instruction

Meta-analyses included in review

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18. https://doi.org/10.1037/a0021017

Baker, S., Gersten, R., & Lee, D-S. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *Elementary School Journal*, *103*(1), 51–73. https://doi.org/10.1086/499715

Bangert-Drowns, R. L., & Bankert, E. (1990). Meta-analysis of effects of explicit instruction for critical thinking.

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review*, *30*(3), 973–999. https://doi.org/10.1007/s10648-018-9436-8

Dennis, M. S., Sharp, E., Chovanes, J., Thomas, A., Burns, R. M., Custer, B., & Park, J. (2016). A meta-analysis of empirical research on teaching students with mathematics learning difficulties. *Learning Disabilities Research & Practice*, *31*(3), 156–168. https://doi.org/10.1111/ldrp.12107

Gersten, R., & Baker, S. (2001). Teaching expressive writing to students with learning disabilities: A meta-analysis. *Elementary School Journal*, *101*(3), 551–572. https://doi.org/10.1086/499668

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*(3), 1202–1242. https://doi.org/10.3102/0034654309334431

Gillespie, A., & Graham, S. (2014). A meta-analysis of writing interventions for students with learning disabilities. *Exceptional Children*, *80*(4), 454–473. https://doi.org/10.1177/0014402914527238

Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *NASSP Bulletin*, *89*(642), 24–46. https://doi.org/10.1177/019263650508964204

Jacobse, A. E., & Harskamp, E. (2011). *A meta-analysis of the effects of instructional interventions on students' mathematics achievement*. GION, Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling, Rijksuniversiteit Groningen.

Stockard, J., Wood, T. W., Coughlin, C., & Rasplica Khoury, C. (2018). The effectiveness of Direct Instruction curricula: A meta-analysis of a half century of research. *Review of Educational Research*, *88*(4), 479–507. https://doi.org/10.3102/0034654317751919

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, *68*(3), 277–321. https://doi.org/10.3102/00346543068003277

Other references cited

Archer, A. L., & Hughes, C. A. (2011). *Explicit instruction: Effective and efficient teaching*. Guilford Press.

Engelmann, S. (1967). Relationship between psychological theories and the act of teaching. *Journal of School Psychology*, *5*(2), 93–100. https://doi.org/10.1016/0022-4405(67)90022-2

Engelmann, S. (2007). *Teaching needy kids in our backward system: 42 years of trying*. ADI Press.

Klahr, D. (2009). "To every thing there is a season, and a time to every purpose under the heavens": What about direct instruction? In S. Tobias & T. M. Duffy (Eds.), *Constructivist instruction: Success or failure?* (pp. 291–310). Routledge.

Mastery learning

Meta-analyses included in review

Arhin, D., & Opoku, W. B. (2000). A meta-analysis on effects of mastery learning strategy (MLS) on academic achievements of learners. *International Journal of Research and Innovation in Social Science, 4*(7), 102–106.

Block, J. H., & Burns, R. B. (1976). Mastery learning. *Review of Research in Education*, *4*, 3–49.

Education Endowment Foundation. (2021). *Teaching and learning toolkit: Mastery learning*. https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit

Guskey, T. R., & Gates, S. L. (1986). Synthesis of research on the effects of mastery learning in elementary and secondary classrooms. *Educational Leadership*, *43*, 73–80.

Guskey, T. R., & Pigott, T. D. (1988). Research on group-based mastery learning programs: A meta-analysis. *Journal of Educational Research*, *81*(4), 197–216. https://doi.org/10.1080/00220671.1988.10885824

Hodgen, J., Foster, C., Marks, R., & Brown, M. (2018). *Improving mathematics in key stages two and three: Evidence review*. Education Endowment Foundation. https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/maths-ks-2-3

Kulik, C. C., & Kulik, J. A. (1986). Mastery testing and student learning: A meta-analysis. *Journal of Educational Technology Systems*, *15*(3), 325–345. https://doi.org/10.2190/FG7X-7Q9V-JX8M-RDJP

Kulik, C. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis: with discussion. *Review of Educational Research*, *60*, 265–307. https://doi.org/10.3102/00346543060002265

Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of instructional improvement in algebra: A systematic review and meta-analysis. *Review of Educational Research*, *80*(3), 372–400. https://doi.org/10.3102/0034654310374880

Slavin, R. E. (1990). Mastery learning re-reconsidered. *Review of Educational Research*, *60*(2), 300–302. https://doi.org/10.3102/00346543060002300

Willett, J. B., Yamashita, J. J. M., & Anderson, R. D. (1983). A meta-analysis of instructional systems applied in science teaching. *Journal of Research in Science Teaching*, *20*(5), 405–417. https://doi.org/10.1002/tea.3660200505

Other references cited

Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment, 1*(2).

Keller, F. S. (1968). Good-bye, teacher.... *Journal of Applied Behavior Analysis*, *1*(1), 79–89. https://doi.org/10.1901/jaba.1968.1-79


Australian
Education
Research
Organisation