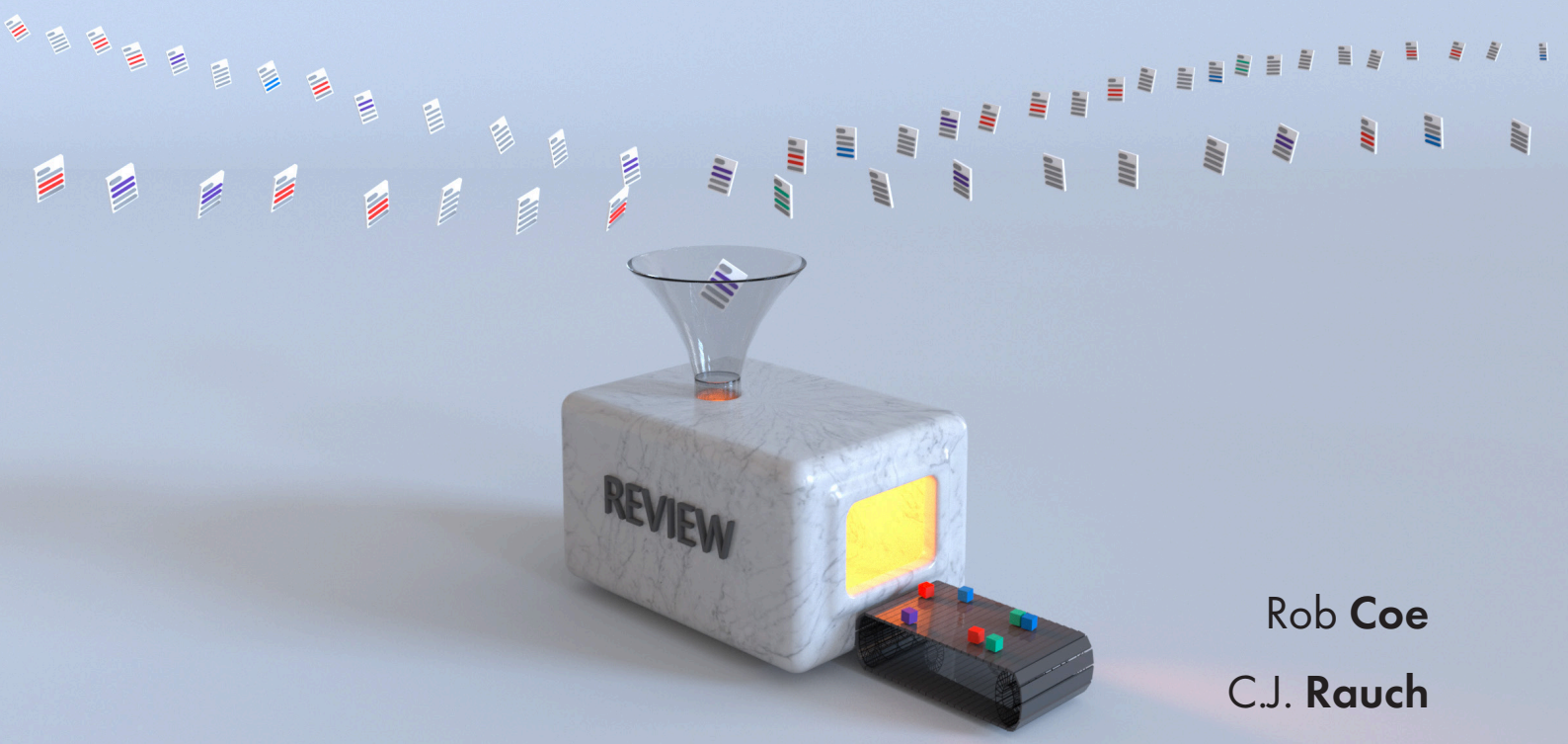
Evidence Based Education

Great Teaching Toolkit

Evidence Review

June 2020



Rob **Coe**

C.J. **Rauch**

Stuart **Kime**

Dan **Singleton**

In partnership with

Cambridge Assessment International Education

greatteaching.com

Acknowledgements

Without collaboration, generosity and a shared will to help every teacher improve, there would be no Great Teaching Toolkit Evidence Review. The work so far has been a truly international project, with 74 collaborators from eleven countries around the world. To all these people listed below – each of whom gave their time and wise feedback freely – we say thank you.

We would also like to take this opportunity to express our thanks to Cambridge Assessment International Education, whose support has helped make this review possible.

Finally, particular credit is reserved for our partners at Ignio, with whom we work closely on many of our developments. The tireless input and invaluable expertise of Dan Singleton and his team is often unseen, but without it, you would not be reading this document today.

As you will discover, the Evidence Review is only the start of a project to transform teacher development. We look forward to acknowledging the collaboration and wisdom of many more teachers, leaders, researchers, designers and policy-makers as we take our next steps together.

Thank you to:

Janice **Allen***
Falinge Park High School, Rochdale

Fatima **Altawil**
Bunat Alghad Academy, Jordan

Phil **Anderson**
Teach First

Paul **Beedle**
Cambridge Assessment International Education

Adam **Boxer**
The Totteridge Academy

Hazel **Brinkworth**
Dulwich College Shanghai Pudong

Gerard **Calnin**
Education University of Hong Kong

Ruth **Carney**
Falinge Park High School, Rochdale

Robbie **Coleman**
Education Endowment Foundation

Steve **Dalgarno**
Dulwich College Suzhou

Shaun **Daly**
Teach First

Rob **Davies**
Swansea Council

David **Didau**
The Learning Spy

Megan **Dixon**
Aspire Educational Trust

Anne **Dwyer**
Yew Wah International Education Foundation

Mark **Enser**
Heathfield Community College

Katie **Fas**
Teach First

Sarah **Flaherty**
Teach First

Harry **Fletcher-Wood**
Ambition Institute

Victoria **Foster**
Dulwich College Shanghai Pudong

Deep **Ghataura**
Heston Community School

Sylwia **Glazewska**
Falinge Park High School, Rochdale

Dave **Greenshields**
Grace College, Gateshead

Rory **Gribbell**
Department for Education, England

Regan **Gurung**
Oregan State University

John **Hattie***
University of Melbourne

Roger **Higgins**
Norwich Research School

Claire **Hill**
Turner Free School, Folkestone

Heather **Hill***
Harvard University

Matt **Hood**
Oak National Academy

Toby **Horrocks**
Ambition Institute

Kevin **House**
Dulwich College International

Kat **Howard**
The Duston School, Northampton

Nicolas **Hübner**
Eberhard Karls University of Tübingen

Ben **Hughes**
Teach First

Ahmed **Hussain**
Wellington College International

Edward **James**
Ambition Institute

Rory **Johnson**
Falinge Park High School, Rochdale

Mark **Jones**
Dulwich College Suzhou

Niki **Kaiser**
Norwich Research School

Ian **Kelleher**
Center for Transformative Teaching and Learning

Victoria **Kelly**
Teach First

Max **Knogler**
Technical University of Munich

Carolina **Kuepper-Tetzel***
University of Glasgow

Peps **Mccrea**
Ambition Institute

Steve **Mills**
Whitehill Junior School

Rachel **Miller**
Teach First

Rebecca **Mitchell**
Teach First

Sophie **Morgan-Williams**
Ambition Institute

Lindsay **Morlock**
South Shields School

Ross **Morrison McGill**
Teacher Toolkit

Daniel **Muijs**
Ofsted

Ryan **O'Kane**
Falinge Park High School, Rochdale

Paula **O'Reilly**
Falinge Park High School, Rochdale

Rachel **Ogden**
Falinge Park High School, Rochdale

Rachel **Orr**
Teach First

Katy **Pautz**
Falinge Park High School, Rochdale

Lisa **Pettifer**
Trinity School, Carlisle

Morgan **Polikoff***
University of Southern California

Alex **Quigley**
Alex Quigley Consults

JoAn **Radojkovich**
Dulwich College International

Nick **Rose**
Ambition Institute

Amanda **Seward**
Teach First

Tom **Sherrington**
Teacherhead

Karen **Taylor**
International School of Geneva

Natasha **Tyrwhitt-Drake**
Teach First

Adrie **Visscher**
University of Twente

Mick **Walker****
Chartered Institute of Educational Assessors

Paul **Warwick**
University of Cambridge

Helen Jennifer **Webb**
Teach First

David **Weston**
Teacher Development Trust

Dylan **Wiliam***
UCL Institute of Education

James **de Winter**
University of Cambridge

Edward **Wright***
Eltham College

* EBE Advisory Board

**EBE Advisory Board Chair

© 2020 Evidence Based Education

Published by Evidence Based Education
in partnership with
Cambridge Assessment International Education

Designed by
IGNIO

Foreword

Dr Tristian Stobie

Director, Curriculum and Qualifications Development

Cambridge Assessment International Education

Teaching should be a rewarding profession where teachers are empowered and supported to be the best creative professionals they can be. The overwhelming body of research finds that the most important factor in improving student outcomes is good teaching. Therefore, helping teachers become better is the most important responsibility we have as educational leaders, as it is the best way to help learners fulfil their potential.

Unfortunately, teacher autonomy, creativity and trust have been eroded in recent decades in some educational systems, by a drive toward compliance. While the goal has been noble – to measure and rank institutional and individual performance, increase accountability and reduce variability – the unintended consequence has often been to reduce teacher learning to formulaic practice. In these settings, feedback to teachers has not been as supportive and informative as it should be to give them control and ownership over their professional development and practice.

In contrast, the Great Teaching Toolkit is a breath of fresh air – treating teachers like the professionals they are. It provides both a synthesis of evidence from authoritative studies, and the findings of this evidence, that teachers can relate to their own experience. What makes it so valuable is its clear focus on areas of practice that have the potential to improve student learning and outcomes.

Professional learning happens when we think hard about our practice and take full ownership of it. Cambridge International is pleased to be able to sponsor this review, which clearly defines what is worth teachers thinking hard about. These are principles and practices that we endorse and use in developing our own professional development services to schools, with the aim of helping teachers become confident, responsible, reflective, innovative and engaged

At the time of writing, the educational world is in turmoil caused by the Covid-19 crisis. Teachers have had to learn quickly to adapt, teach online and support learners in new ways. A number of commentators have speculated on the implications for the future of schools and the nature of the teaching profession. In such a climate, the evidence-based insights provided in the Great Teaching Toolkit are even more significant. We believe that the Toolkit's universal and timely principles will be an invaluable resource to teachers and schools around the world.

Executive Summary

What are the best bets for teachers to invest time and effort in if they want their students to learn more?

We have reviewed existing research studies and frameworks that are relevant to the components and routes to improvement of teacher effectiveness. Our aim is to help teachers make better decisions about what they can best do to improve their effectiveness. In summary, we have identified four priorities for teachers who want to help their student learn more:

1. understand the content they are teaching and how it is learnt
2. create a supportive environment for learning
3. manage the classroom to maximise the opportunity to learn
4. present content, activities and interactions that activate their students' thinking

We present a model that comprises these four overarching dimensions, with a total of 17 elements within them. An 'element' is defined as something that may be worth investing time and effort to work on to build a specific competency, skill or knowledge, or to enhance the learning environment. There is no implication that the complexity of teaching can be reduced to a set of techniques, but evidence suggests the best route to expertise is likely to involve a focus on developing competencies, guided by formative feedback in a supportive professional learning environment.

This review is the first stage of an ambitious wider project to create a 'Toolkit' that will:

- personalise the curriculum for teacher learning (according to ages and subjects taught, school context and student characteristics, current profile of expertise, etc.)
- develop systems and instruments to provide formative, actionable feedback that helps teachers to focus their learning, evaluate their impact and track their professional growth
- coordinate networks for peer and expert support to generate, share and apply evidence about the most effective ways to improve

The individual elements of the model for Great Teaching are as follows.

1. Understanding the content

1 Having deep and fluent knowledge and flexible understanding of the content you are teaching

2 Knowledge of the requirements of curriculum sequencing and dependencies in relation to the content and ideas you are teaching

3 Knowledge of relevant curriculum tasks, assessments and activities, their diagnostic and didactic potential; being able to generate varied explanations and multiple representations/analogies/examples for the ideas you are teaching

4 Knowledge of common student strategies, misconceptions and sticking points in relation to the content you are teaching

2. Creating a supportive environment

1 Promoting interactions and relationships with all students that are based on mutual respect, care, empathy and warmth; avoiding negative emotions in interactions with students; being sensitive to the individual needs, emotions, culture and beliefs of students

2 Promoting a positive climate of student-student relationships, characterised by respect, trust, cooperation and care

3 Promoting learner motivation through feelings of competence, autonomy and relatedness

4 Creating a climate of high expectations, with high challenge and high trust, so learners feel it is okay to have a go; encouraging learners to attribute their success or failure to things they can change

3. Maximising opportunity to learn

1 Managing time and resources efficiently in the classroom to maximise productivity and minimise wasted time (e.g., starts, transitions); giving clear instructions so students understand what they should be doing; using (and explicitly teaching) routines to make transitions smooth

2 Ensuring that rules, expectations and consequences for behaviour are explicit, clear and consistently applied

3 Preventing, anticipating & responding to potentially disruptive incidents; reinforcing positive student behaviours; signalling awareness of what is happening in the classroom and responding appropriately

4. Activating hard thinking

1 Structuring: giving students an appropriate sequence of learning tasks; signalling learning objectives, rationale, overview, key ideas and stages of progress; matching tasks to learners' needs and readiness; scaffolding and supporting to make tasks accessible to all, but gradually removed so that all students succeed at the required level

2 Explaining: presenting and communicating new ideas clearly, with concise, appropriate, engaging explanations; connecting new ideas to what has previously been learnt (and re-activating/checking that prior knowledge); using examples (and non-examples) appropriately to help learners understand and build connections; modelling/demonstrating new skills or procedures with appropriate scaffolding and challenge; using worked/part-worked examples

3 Questioning: using questions and dialogue to promote elaboration and connected, flexible thinking among learners (e.g., 'Why?', 'Compare', etc.); using questions to elicit student thinking; getting responses from all students; using high-quality assessment to evidence learning; interpreting, communicating and responding to assessment evidence appropriately

4 Interacting: responding appropriately to feedback from students about their thinking/knowledge/understanding; giving students actionable feedback to guide their learning

5 Embedding: giving students tasks that embed and reinforce learning; requiring them to practise until learning is fluent and secure; ensuring that once-learnt material is reviewed/revisited to prevent forgetting

6 Activating: helping students to plan, regulate and monitor their own learning; progressing appropriately from structured to more independent learning as students develop knowledge and expertise

Contents

The Great Teaching Toolkit

The fundamental goal of everyone that works in education is to improve students' lives. While many personal, family, and cultural factors contribute to students' outcomes, a large body of research indicates that what teachers do, know and believe matters more to the achievement of students than anything else we can influence. The quality of teaching is hugely important to the outcomes of young people, and great teaching can be learnt. Raising the quality of teaching within existing schools is probably the single most effective thing we could do to promote both overall attainment and equity (Wiliam, 2018).

Teachers know a lot about learning and how to make it happen. Unfortunately, it seems common for that knowledge to be disregarded when it comes to their own **professional learning**. Among the conditions we would routinely provide for our students are a clear and sequenced curriculum that sets out the learning aims, diagnostic assessment to ensure prerequisites are secure, models of excellent performance, scaffolding, guidance, opportunities for practice and, crucially, feedback that guides next steps and indicates progress. Many teachers who strive to ensure their students' learning has all these supports would say that their own has none of them.

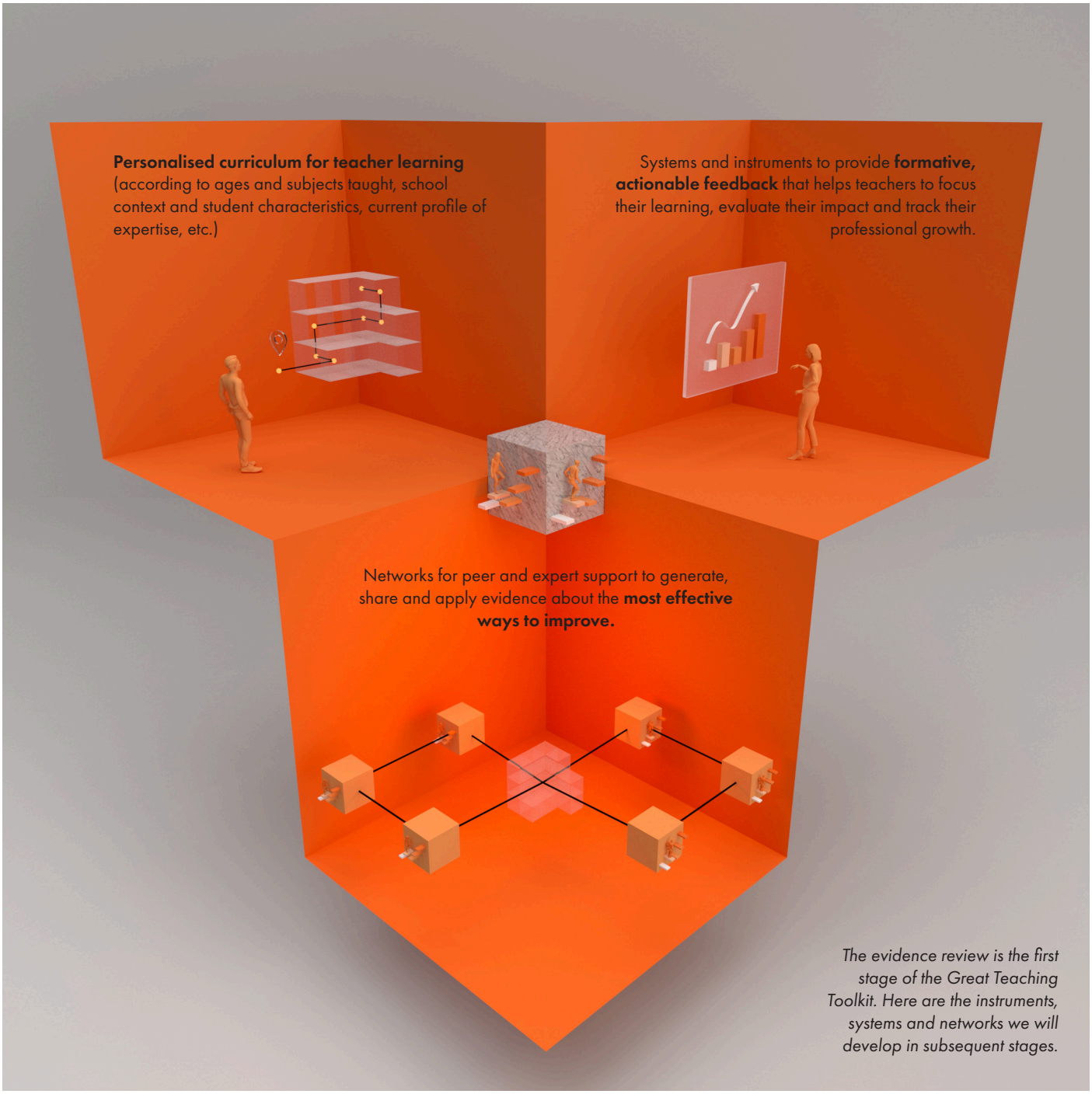
Professional learning:
Teachers' professional learning continues beyond their initial teacher education. While "inset days" or "twilight sessions" may be what first springs to mind, it also can include mentorship, engagement with research, deepening knowledge of the content, or any other activities that aim to improve teachers' effectiveness.

Fortunately, human beings can get really good at quite complex tasks if they just have good feedback that tells them whether they are succeeding. Unfortunately, the kinds of feedback that teachers can easily get about their classroom practice are often not very helpful. Creating feedback systems that enable continuous improvement is an area of focus we committed to in our 2019 **Manifesto**, which outlines what we believe an evidence-informed education system should look like. Systems with good feedback can become self-improving as participants learn to optimise outcomes – students benefit directly from this. But when feedback is seen as supportive it can also have real benefits for teachers, giving them agency and control over their professional development and satisfaction and engagement in the process (Coe, 1998), and subsequently for school and system leaders.

How will we create a feedback system to better enable effective teaching? The Great Teaching Toolkit is how, and it starts with this report. We can think of it as a model for teacher learning. It gives us a credible summary of the elements of great teaching practice, the kind that impacts most on learning. Following this report, we will develop and release a set of instruments to help teachers anonymously assess their strengths and identify their own development priorities in the areas identified in this report. The same tools will provide diagnostic formative feedback for teachers as they work on specific goals to improve their practice. Although teaching is an extremely complex set of practices and definitely not just a set of techniques or recipes, taking a

specific technique, skill or area of knowledge and practising to a high level of proficiency is a key way to improve overall effectiveness.

The Great Teaching Toolkit will also aim to identify the kind of professional development that leads to improvement in specific areas of practice. This stage of the project will require a community of thousands of educators working toward a shared aim, supporting each other and creating the evidence we need. The strong, overarching goal here is to help teachers take ownership of their professional learning and to help them enhance their practice for the benefit of students.

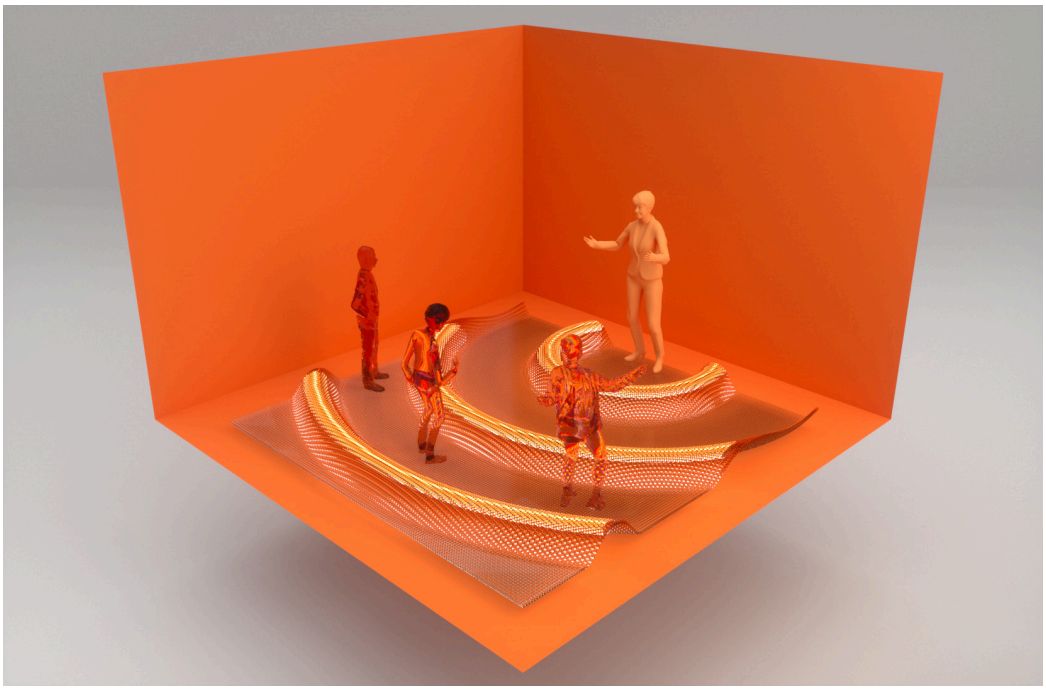


The evidence review is the first stage of the Great Teaching Toolkit. Here are the instruments, systems and networks we will develop in subsequent stages.

Great teaching must be defined by its impact: a great teacher is one whose students learn more. It cannot be defined by compliance to a particular set of practices, however soundly based, nor by the demonstration of specific skills – nor, even, by the possession of particular teacher mindsets or understandings. Teaching is complex.

However, the evidence we present here makes it clear that, on balance, having these things is better than not having them. We also have good evidence that engaging in systematic, focused efforts to develop fluency and expertise in these skills and practices, and to develop teachers' understanding of the principles and theory underpinning them, are likely to be our best bets for enhancing impact. And none of this happens in isolation: great teachers have a drive to improve their impact and to collaborate with and support their colleagues to improve.

Everyone in every walk of life can be better. Every teacher, no matter how experienced, can improve, if they want to and have the support to. But, as a teacher, even when you decide to take that step, it's often difficult to know where to start. Your resources are precious, you have no time to waste. How should you prioritise your professional development? What are your best bets in terms of making the most difference to your students? We hope this review, and the rest of the Great Teaching Toolkit, will help to answer those questions.



Great teaching must be defined by its impact

Review of the evidence

We set out to identify, review and then summarise the best available evidence (drawn from both academic reviews and existing frameworks) about what teachers' practices, skills, knowledge and behaviours are important for students' learning. We did this to ascertain what the evidence suggests is important for teachers' learning.

In addition, we set out to review the related evidence on measuring these important features, and to identify useful **indicators** of things found to be associated with student learning (a supportive teacher-student relationship, for instance) that might help us provide better feedback for teachers' professional learning. An important step for the future development of simple, powerful tools is to help teachers truly understand the evidence in a way that would make it actionable: to bring it to life and operationalise it.

Indicators:
When working with concepts that are difficult to measure directly, it is possible to use an indicator as a way to make conclusions about the topic. A more valid indicator allows more valid conclusions. For example, it is common to look at student exam results as an indicator of attainment. In another example later, teacher behaviours are an indicator for the complex concept of activating thinking.

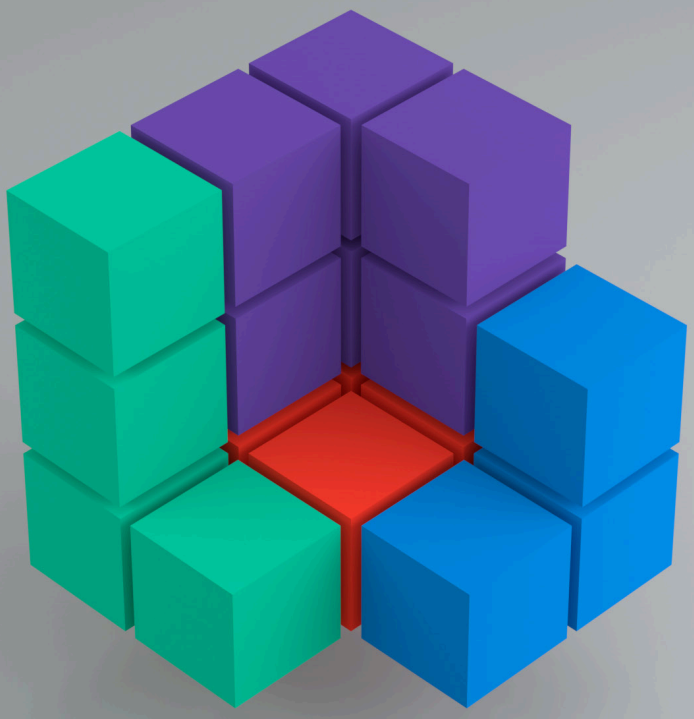
What we found was a consensus within the existing research – a signal within the noise – about which elements of teaching appear to be worth learning. Simultaneously, we also found that the evidence base is limited; for example, there is a predominance of **correlational studies** over those making strong causal claims, something we say more about later.

Correlational studies:
Much of the available research is based around correlational studies; in these the relationships between two variables is measured. While interesting, the conclusions drawn from them are limited. We cannot tell if the two have a causal relationship – does X cause Y, or does Y cause X? Or might there be a third variable, Z? Therefore, while we may find a postive correlation between a teaching practice and student outcomes, we do not know if the practice caused the outcome.

Limitations such as this will make developing certain aspects of the Great Teaching Toolkit very challenging, but we believe that – together – we can overcome these challenges. We believe they necessitate a new collaboration between classroom practitioners, academic researchers, designers and innovators; one which develops and tests a model for Great Teaching and delivers feedback tools that help teachers know where they are, where they're heading, and how to get there.

What follows, then, is the starting point: a simple, digestible summary of what a large and complex body of evidence says about what is worth learning.

A Model for Great Teaching



A Model for Great Teaching

Rationale for presenting a model

In an ideal world, we would already have a conceptually clear and empirically well-validated model of classroom teaching that would make it explicit what great teaching looks like and how to get more of it. The model would take account of differences in the ages and other characteristics of the learners. It would factor in the subjects – or even topics – being taught, and relevant features of the context or school. We would also have a curriculum model for teachers' professional learning that set out what teachers need to learn to become better teachers, according to their current profile of strengths and weaknesses and the context in which they work. Such a **curriculum** would be **sequenced** and prioritised: prerequisites and dependencies would be known and clearly set out; the likely 'payback', in terms of increases in student attainment, for each hour spent on particular teacher development activities, would be quantified and optimised.

Curriculum sequencing:
Mastery of certain content may require understanding of certain prior knowledge. Sequencing identifies these prerequisites within the curriculum so they can be taught and assessed in an appropriate, logical order. For example, a learner's mastery of π depends on their understanding of diameter, radius, and circumference; the sequencing should identify this.

Unfortunately, we do not currently have either of these things. Instead of a clear, comprehensive and reliable model of great teaching, research gives us partial insights, often contradictory or confusing, much of it based on weak correlations between ill-defined teacher behaviours and rather impoverished measures of student learning that may reflect confounds as much as genuine **causal relationships**. Where we have stronger causal designs – the kind that might allow us to infer that training or development for teachers in particular competences leads to enhanced student learning – the results have often been inconsistent or disappointing. And instead of a well-specified curriculum for teacher learning, we have lots of traditions and loud claims, whose projected confidence or popularity seems to outweigh their evidential warrant, and whose relative merits are hard to evaluate.

Causal relationship:
A relationship in which it has been shown, usually through a controlled experiment, that one variable (independent) causes the other (dependent)

One insight we do have is that these two are not the same thing. Being able to describe great teaching is not the same as knowing how to get more of it. Our interest is more in the latter: knowing what great teachers should do to become even greater, or how teachers who are not as great as they could be could become great.

This leads us to what might at first sight seem like a rather narrow and reductionist project, breaking down a complex, nuanced, beautiful thing like 'great teaching' into an atomised list of competences. But this is familiar territory for anyone who has tried to become expert in any complex activity or performance, whether in sport, music, dance, writing, art – or professionals such as pilots, doctors, lawyers or teachers. Giving a precise and useful definition of great performance may be impossible but, despite that, we generally do know something about the steps that lead to expertise. And this usually means breaking the complex activity down into components and exercises, clarifying, then practising them with appropriate guidance until they

are fluent and proficient, and integrating those isolated techniques back into the complex and mysterious whole (Ericsson, 2009).

Our starting point for this 'curriculum' is to identify the elements of great teaching that come out of existing research and then to investigate the process of trying to get better at each of them in isolation. This does not imply that we think classroom teaching can be reduced to a set of isolated techniques; only that our best bet for learning to be a better teacher is to work on specific, underpinning competences, one at a time. We are likely to find that some can be improved more quickly than others; that some matter more than others in their impact on student learning; that there are interactions, dependencies and threshold effects in their relationships; that priorities should be different for different teachers at different stages, in different contexts. As we discover and incorporate these complexities, we hope our model will become more useful.

Our aim is to help teachers make better decisions about what they can best do to improve their effectiveness. We know that, as with other kinds of learning, teachers' professional learning is most effective when the content and activities are targeted to be appropriate to the needs and existing capabilities of the learner (Creemers et al., 2013). It follows that the answer to the question 'What can I best focus on to improve?' is likely to be different for different teachers. We hope that our model can be used to help teachers make more evidence-based, individualised decisions about how to spend a limited amount of time for professional development to get the biggest return in enhanced student learning.

Overview: The Elements of Great Teaching

Our model for great teaching presents a simple narrative:

Great teachers:

1. understand the content they are teaching and how it is learnt
2. create a supportive environment for learning
3. manage the classroom to maximise opportunity to learn
4. present content, activities and interactions that activate their students' thinking

For each of these four broad dimensions, we break it down into a set of elements. An 'element' here is defined as something that may be worth investing time and effort to work on. It may capture a specific skill, technique or area of knowledge that great teachers appear to have: what we have called a 'competency'. But in some cases, the element may be more an environmental than a behavioural indicator. For example, indicators of classroom climate or relationships may not point to a particular teacher behaviour or competency but may still capture an aspect of great teaching. The precise behaviours or actions a teacher should do are not specified, but the objectives and success criteria for their learning are clear. We also recognise that the word 'competency' carries some unfortunate baggage in certain contexts, either being associated with competency-based frameworks in accountability models, or denoting over-generalised skills that are supposedly transferable across domains; neither is part of our intended meaning.

At this stage, there is a degree of arbitrariness to the model. The four dimensions overlap in some areas and their boundaries are debatable. Most of the elements could be further split into smaller strands, which might be conceptually purer and make it easier to practise or learn to improve them; this would also multiply the complexity of the model. We have to start somewhere, but fully expect some of these decisions to be revised as we get more experience of working with the model.

A further challenge is the tension between wanting a generic model, that captures some universal principles of great teaching, and acknowledging that the manifestations of great teaching across ages, contexts and subjects appear very diverse. We think the generic principles are useful and important (and supported by evidence), partly because great teachers need to understand the principles of how and why different techniques are effective and when to deploy them. Nevertheless, it is important to remember that most of these elements will look very different in different classrooms, and their relative importance will also vary.

With these caveats in mind, we offer an overview of each dimension and a more detailed, practice-focused description of its different elements, what exactly each one means and the evidence behind it.

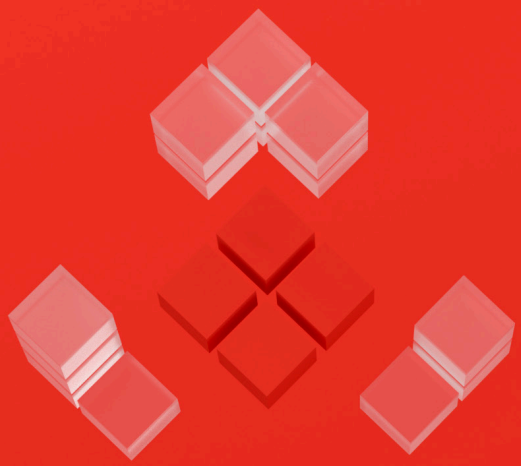
Evidence for a four-dimensional model

There is no universal consensus in the research literature about how many dimensions a research-based model of teaching should contain. We find the arguments set out by Praetorius et al. (2018) compelling, that their three-dimensional model captures a reasonable consensus of evidence from a range of existing studies, though even their own evidence does not seem to support it unequivocally. Certainly, other frameworks present it differently. For example, the Dynamic Model (Creemers and Kyriakides, 2011) has eight classroom dimensions, ISTOF (Muijs et al, 2018) has seven components, ICALT (van de Grift et al. 2017) has six, Rosenshine (2010) has ten principles, and the Early Career Framework has eight standards. However, their content is readily compatible with the aforementioned three-dimensional model; ultimately, it seems to be a somewhat arbitrary choice. Moreover, the three-dimensional model lends itself easily to a simple narrative about what great teachers do: they create a supportive environment for learning, they maximise opportunity to learn and they activate their students' thinking.

However, we have also been convinced by the arguments that a fourth dimension should be included: content knowledge. This is missing from the generic models that focus on observable classroom behaviours, for obvious reasons: it is more about teacher knowledge than teacher behaviour. We recognise that there is a danger here – there is no point in teachers having good content knowledge if their classroom actions do not reflect this. Indeed, in some of the observational frameworks, content knowledge is included in that way. But there is enough evidence that effective teachers need to have particular kinds of knowledge and understanding of the material they are teaching to justify including it here as something that some teachers could profitably work on. Because it is likely to be a prerequisite rather than an extended focus of professional learning, we place this first.

01

Understanding the content



Great teachers understand the content they are teaching and how it is learnt

This means teachers should have deep and fluent knowledge and flexible understanding of the content they are teaching and how it is learnt, including its inherent dependencies. They should have an explicit repertoire of well-crafted explanations, examples and tasks for each topic they teach.

Summary of Dimension 1

1.1 Having deep and fluent knowledge and flexible understanding of the content you are teaching

1.2 Knowledge of the requirements of curriculum sequencing and dependencies in relation to the content and ideas you are teaching

1.3 Knowledge of relevant curriculum tasks, assessments and activities, their diagnostic and didactic potential; being able to generate varied explanations and multiple representations/analogies/examples for the ideas you are teaching

1.4 Knowledge of common student strategies, misconceptions and sticking points in relation to the content you are teaching

Elements of Dimension 1

1

The first element of Dimension 1 is essentially content knowledge, of a deep and connected kind. Teachers need to know how different ideas in the subject or domain are related, similar, sequential, analogous or distinct. They need to have thought about, and have good answers to, the kinds of 'Why?' and 'What would happen if...?' questions that students may ask and that teachers themselves should ask to promote connected and higher-order thinking. They should be able to solve the kinds of problems they must help students to solve, and to produce model answers that exhibit the skills and knowledge they need their students to learn, without errors. We might also include, under the heading of **content knowledge**, teachers' theoretical knowledge of the domain of learning. An example would be the requirement for teachers of reading to understand morphology, "the ways in which morphemes communicate meaning and govern spelling construction" (Castles et al., 2018). This requires more than just being able to read well themselves, but also to know about the fundamental anatomy of the reading process.

Content knowledge:
A teacher's knowledge and understanding of the subject(s)

2

A second aspect moves us from what is usually classified as 'content knowledge' (CK) to '**pedagogical content knowledge**' (PCK). This distinction was originally made by Shulman (1986; see also Ball et al., 2008), though a range of different interpretations of PCK have since been offered. This aspect of PCK involves knowing and being able to explain the dependencies and connections among different parts of the curriculum, and hence the requirements for sequencing. If you want students to learn a specific topic, what knowledge and skills must they have already to enable this new learning? If a student is struggling with a particular idea or technique, what kinds of gaps in underpinning knowledge might be the explanation? For each new idea, what connections do learners need to make with previous knowledge? This kind of teacher curriculum knowledge is exemplified in curriculum planning, schemes of work and lesson plans that depend on correct sequencing and planned reactivation of prior knowledge.

Pedagogical content knowledge:
While it has various nuanced definitions, the key idea to pedagogical content knowledge is that it is more than just knowledge about the content itself, but the learning associated with that particular content. PCK and content knowledge are included in separate elements, emphasing the difference between the two.

3

The third element of this dimension is knowledge of curriculum tasks and activities, and of standard explanations, models, analogies, representations and examples to explain and convey hard ideas. Expertise in teaching a particular topic requires having a repertoire of appropriate activities, but in particular, understanding "the **didactic** and diagnostic potential of tasks, their cognitive demands and the prior knowledge they implicitly require" (Baumert & Kunter, 2013). Expert teachers are readily able to generate or select learning activities that are appropriate for the level of challenge required or that elicit diagnostic information about learners' thinking. As with all these elements of content knowledge, this expertise is likely to be very topic-specific: the same geography teacher may be easily able to identify great resources for teaching map skills, but have a much less rich repertoire for glaciation, for example.

Didactic:
A didactic task is one in which information is explicitly transferred to a learner.

For each topic they teach, great teachers will have learnt effective ways of presenting the ideas: explanations that students get. In the classic **direct instruction** model (Adams & Engelmann, 1996), for example, these explanations are carefully refined and scripted, on the grounds that an individual teacher's own spontaneous explanation is unlikely to be as good as a high-quality scripted presentation.

Direct instruction:
Direct instruction has taken on many meanings. In this particular example, it refers to a particular programme of specific, generally scripted, practices.

In presenting abstract ideas, great teachers use analogies, models and representations to help learners visualise the concepts and relate them to what they already know. For example, the ball and stick model in chemistry represents molecules in a concrete, visual way that facilitates understanding of why atoms bond in particular ways. It is an effective way to introduce the ideas, but of course is not actually true, and has to be revised as students' understanding becomes more advanced. Another example would be the use of manipulatives and representations in teaching early mathematics (EEF, 2020), which can be effective in helping children to engage with and understand abstract ideas about number. Selecting good examples and non-examples (e.g., using the Frayer Model[1]) is another way of making new vocabulary or abstract ideas concrete. However, even with the best explanation, some students still may not get it. Teachers need to have more than one way of explaining or presenting the idea, and multiple examples and non-examples (ideally tailored to the student's particular misconception or gap), so that they can keep going until the student does get it.

The key point about these explanations, models, analogies, representations and examples is that they form part of the teacher's pedagogical content knowledge. In many systems, teachers are expected to learn these on the job, through trial and error, experience, intuition and *ad hoc* sharing. But this knowledge can also be explicitly taught. Great teachers also have access to great materials, rather than being expected to search for or create their own.[2]

1 For example, see Alex Quigley's blog on using the Frayer Model to teach vocabulary: **https://www.theconfidentteacher.com/2018/04/vocabulary-knowledge-and-the-frayer-model/**

2 An example from the US is **edreports.org**, which provides evidence-based reviews of textbooks and instructional materials.

4

Our fourth and final element is a knowledge of student thinking and, in particular, the misconceptions, typical errors and types of strategies students exhibit. Student misconceptions around particular ideas are predictable and inevitable. Great teachers design their presentations and learning activities to anticipate and address these misconceptions directly and explicitly, both by exposing and challenging the misconception and by presenting the correct conception clearly and directly.

A final point to note for all these aspects of teachers' understanding of curriculum content is that they are very much necessary but not sufficient for effective practice. Knowing students' likely misconceptions has no benefit unless lessons and delivery are structured to address them; having a repertoire of good examples is only useful if they are employed appropriately. In general, pedagogical content knowledge (PCK) should be learnt and deployed in the context of classroom practice: theoretical knowledge alone is not enough. This may be one of the reasons that evaluations of the impact on student learning of attempts to increase teachers' PCK have sometimes had disappointing results. It is certainly possible that we could have placed some of these elements in Dimension 4, which is concerned with teachers' classroom practices to activate student thinking: for example, 'having multiple explanations, examples, etc.' has considerable overlap with 'explaining' (Element 2 of Dimension 4, below) which is about actually using these explanations and examples effectively.

Evidence for Dimension 1

The evidence for the importance of 'pure' content knowledge is a bit mixed and conceptually somewhat confused. Many studies that have looked for relationships between teachers' qualifications or advanced subject knowledge and learning gains have failed to find them consistently (Wayne & Youngs, 2003). Nevertheless, plenty of studies have shown that measures of teachers' knowledge and conceptual understanding of the specific content they are teaching do have some predictive power for their students' learning (Baumert et al., 2010; Hill et al., 2005; Hill & Charalambous, 2012; Lynch et al., 2019; Sadler et al., 2013). These relationships are generally modest-to-weak, probably non-linear and the existing evidence may be limited to particular topics, ages or subjects. For example, Hill et al. (2005) found that variation at the bottom end of their scale of 'Content Knowledge for Teaching' (CKT) was related to effectiveness, but for the majority of teachers, whose content knowledge was at least adequate, there was no further benefit in increased CKT. There is also some evidence that training programmes designed to enhance teachers' content knowledge can lead to enhanced student learning, though again the findings are mixed (Baumert et al., 2010; Lynch et al., 2019; Timperley et al., 2007). Many of the available studies have used mathematics content, so the generalisability to other subjects is unclear, though Kaiser and König (2019) give examples of evidence from other subjects. Metzler and Woessmann (2012) provide evidence of the importance of subject knowledge for Y6 teachers in Peru.

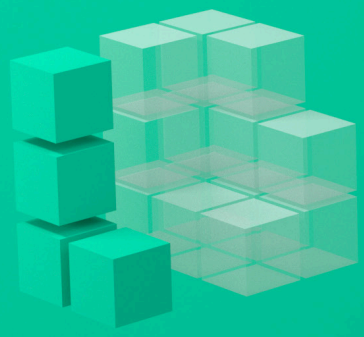
There is broad support for the role of teachers' PCK (see Baumert et al., 2010; Kaiser & König, 2019 for reviews) though, again, much of it is from mathematics and science, and different studies operationalise PCK in different ways. A framework that specifically identifies curriculum and lesson planning-related PCK, and provides evidence of its importance, comes from the TEDS-M project (Teacher Education and Development Study in Mathematics, Blömeke et al., 2016).

"Knowledge of the didactic and diagnostic potential of tasks" is a key component of the COACTIV model of mathematics PCK (Baumert & Kunter, 2013), which was found by Baumert et al. (2010) to be a substantial predictor of student learning, after controlling for a wide range of other variables. The evidence for the importance of teachers' knowledge of good explanations, models, analogies, representations and examples in relation to the content they teach comes from the same sources cited above, for example, Baumert et al. (2010).

Being able to anticipate, identify and address student misconceptions is a feature of a number of models of teaching effectiveness (e.g., Hill et al.'s Mathematical Quality of Instruction or the Early Career Framework for England) and is supported by a range of evidence (e.g., Baumert et al., 2010; Blömeke et al., 2016; Hill et al., 2005; Hill and Chin, 2018). Understanding how 'novice' learners see the world differently from 'experts' has also been claimed as important for teachers (e.g., van Merriënboer et al., 2006), as has an understanding of how 'threshold concepts' – key ideas in a discipline that act as a portal to new ways of thinking and understanding – may either open up new insights or be 'troublesome' barriers (Meyer & Land, 2005). However, direct empirical support for the value of any specific kinds of teacher knowledge about threshold concepts is less clear. Evidence-based approaches to addressing misconceptions include challenging them or simply emphasising the 'scientific' conception (Braasch et al., 2013).

02

Creating a supportive environment



Great teachers create a supportive environment for learning

A supportive environment is characterised by relationships of trust and respect between students and teachers, and among students. It is one in which students are motivated, supported and challenged and have a positive attitude towards their learning.

Summary of Dimension 2

2.1 Promoting interactions and relationships with all students that are based on mutual respect, care, empathy and warmth; avoiding negative emotions in interactions with students; being sensitive to the individual needs, emotions, culture and beliefs of students

2.2 Promoting a positive climate of student-student relationships, characterised by respect, trust, cooperation and care

2.3 Promoting learner motivation through feelings of competence, autonomy and relatedness

2.4 Creating a climate of high expectations, with high challenge and high trust, so learners feel it is okay to have a go; encouraging learners to attribute their success or failure to things they can change

Elements of Dimension 2

1

The first element of this dimension concerns the quality of the relationships between teacher and students. Teachers should show respect and sensitivity towards the individual needs, emotions, culture and beliefs of their students That respect should also be reciprocated: teachers should behave in ways that promote student respect for the integrity and authority of the teacher. Teachers should convey care, empathy and warmth towards their students and avoid negative emotional behaviours, such as using sarcasm, shouting or humiliation. This element is multifaceted and complex, and it is arguable that the range of issues it covers justifies allocating more than one element to it. There are two particular aspects of teacher-student relationships that deserve specific attention: relationships with students with SEND (special educational needs and disabilities) and culturally relevant teaching.

The requirement for respect and sensitivity towards students' individual needs is amplified in both importance and difficulty when those needs are more diverse or extreme. Developing good relationships of trust and respect with students with special educational needs, neurodiversity or disabilities often requires specific knowledge and adaptation. Generic labels such as SEND or their subcategories cover a wide range of individual differences, and the processes by which they become attached to individual students – or may go undiagnosed – are also variable. Great teachers know their students well as individuals, are well informed about the nature and requirements of their students' specific needs and have strategies to accommodate them.

Another key part of this element is the need for teaching to be 'culturally relevant' (Ladson-Billings, 1995): great teachers are aware of, respectful

towards and responsive to the cultural identities of their students. This is particularly important when the students' culture differs from, and has the potential to conflict with, that of the teacher or school. Teachers must ensure that good relationships and academic success are compatible with students honouring their cultural competences, values and identities.

2 The second element views the classroom environment through student-student interactions and relationships. Classrooms where students respect and pay attention to each other's thoughts, and feel safe to express their own thoughts, are more productive for learning. Where students cooperate with each other effectively, they are able to benefit from learning interactions with their peers. By contrast, in classrooms where relationships between students are characterised by aggression, hostility, belittling or disrespect, learning is impeded. The teacher plays a role in promoting these positive student relationships and interactions. This aspect of the classroom environment is an element of the Praetorius et al. (2018) model, the Dynamic Model (Creemers & Kyriakides, 2011) and the CLASS framework (Pianta et al., 2012).

3 The third element of the supportive classroom environment focuses directly on student motivation. Students who are motivated to study, learn, engage and succeed are more likely to do so. In considering motivation, we follow Praetorius et al. (2018) and draw on Deci and Ryan's (2008) self-determination theory (SDT) and, in particular, its application to education (Guay et al., 2008). SDT prioritises the kinds of **motivation** that support the individual's wellbeing and development as much as their task performance. SDT distinguishes between two kinds of motivation: autonomous (which is characterised by a feeling of volition, though may have either intrinsic or extrinsic value that has become part of the individual's identity) and controlled (characterised by feeling "pressure to think, feel, or behave in particular ways", either through explicit, contingent reward/punishment, or "introjected regulation": feelings such as guilt, shame or contingent approval). Autonomous motivation is promoted when individuals feel that three basic needs are met: autonomy, competence and relatedness. Autonomy refers to feeling that they choose their behaviour and that it is aligned with their values and interests. Competence means feeling capable of producing desired outcomes and avoiding undesirable ones. Relatedness means feeling connected with and mutually supported by other people.

Motivation:
In everyday parlance, motivation simply refers to the rationale for an individual's behaviour. Within education, it also refers to a whole field of research with a focus on the complex factors affecting student motivation. As this element demonstrates, there are multiple ways of classifying motivation.

4 The fourth and final element of creating a supportive environment concerns teachers' expectations and attributions. Teachers should demand high standards of work and behaviour from all students, being careful not to convey lower expectations for any subgroup, especially one where a common stereotype may be negative. Even when lower expectations may be indirectly conveyed with good intentions (e.g., praising students for poor work to encourage them; avoiding asking challenging questions to students who seem less confident or helping them sooner when they are stuck), it may still undermine their learning. High expectations may be seen as a form of 'tough

love'. Demanding high standards can mean requiring something that teachers do not genuinely believe is likely, so some suspension of disbelief may be called for. When goals are ambitious and demands are high, learners must feel safe to have a go and take a risk, without feeling pressured or controlled. This requires an environment of trust and a complex balance of asking a lot but still being okay if you get only part of it. And whether students succeed or fail, it matters how they account for it: attributing either success or failure to things they can change (such as how hard they worked or the strategies they used) is more adaptive for future success than attributing results to things that are out of their control (like luck, 'ability', or not having been taught it).

Evidence for Dimension 2

This dimension is one part of the German three-dimensional model (Praetorius et al., 2018) and at the heart of the CLASS framework (Classroom Assessment Scoring System, Pianta et al., 2012). This prominence may partly reflect the origins of CLASS in early years settings, though the development and extension of CLASS to classrooms with older children has shown it is just as important there. Nevertheless, it may be that some aspects of this dimension are more important in some types of classroom setting than others (for example, with younger or more educationally 'at-risk' students, or those for whom schooling is generally a less positive experience; Pianta et al., 2012). Indicators of classroom climate also feature in two internationally validated instruments for measuring teaching quality, ICALT (van de Grift et al., 2017) and ISTOF (Muijs et al., 2018).

The importance of classroom environment and relationships is supported by several prominent psychological theories. Among these are Deci and Ryan's (2008) self-determination theory, which identifies feelings of competence, autonomy and social-relatedness as the requirements for students to be motivated and to achieve. Also invoked are theories of meaningful engagement (Csikszentmihalyi & Schneider, 2000), self-efficacy (Bandura et al., 1996), attachment theory (Bowlby, 1969) and Vygotskian social constructivism (Vygotsky, 1962).

We might justify the need for these positive teacher behaviours on grounds of decency and human rights. But there is also empirical evidence to suggest that they are associated with higher achievement, along with other positive student outcomes (Hamre et al., 2014; Pianta et al., 2012; Praetorius et al., 2018). For example, evaluations of My Teaching Partner (e.g., Allen et al., 2011) show that when teachers work on improving the warmth and supportiveness of classrooms, student outcomes improve. There is also evidence of benefits for attainment from the evaluations of interventions that target social and emotional learning by improving classroom environment (Jones & Doolittle, 2017).

There is evidence that autonomous forms of motivation are more conducive to student attainment, persistence and depth of thinking (Vansteenkiste et al., 2004), though other studies have found mixed results and there may be some confusion in the literature about what kinds of teacher behaviours may be classed as 'autonomy-promoting'. The requirements of autonomy,

competence and relatedness are explicitly observed in the German three-dimensional model (Praetorius et al., 2018). In this framework, supporting autonomy means making work interesting and relevant, avoiding competitiveness or public pressure and allowing students choices about how they work; supporting competence means differentiating the difficulty level of work, adapting the level of support, giving students enough time to think and keep up, and responding positively and constructively to errors; support for social relatedness concerns the relationships between teacher-student and student-student outlined above. Praetorius et al. found an overall small positive association (0.12) between these observed behaviours and student attainment.

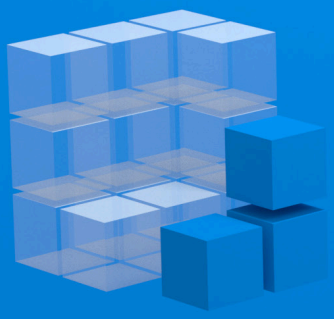
The relationship between high teacher expectations and student attainment has been a mainstay of educational effectiveness research since it began (Muijs et al., 2014). Although much of this research has failed to establish the direction of causality, to conceptualise 'expectations' properly, or to demonstrate that we know how to change teachers' expectations, there probably is enough evidence that both subliminal and explicit teacher expectations can influence student attainment and become, at least to some extent, self-fulfilling prophecies (Muijs et al., 2014). Moreover, it is a characteristic of effective interventions such as mastery learning (Bloom, 1976) that teachers require mastery from all students (Creemers et al., 2013). Another source of theoretical support for high expectations comes from goal-setting theory (Locke & Latham, 2002) which finds that, other things being equal (goals must be specific, accepted, possible and not conflicted), the more challenging the goal, the better the level of performance actually achieved.

Research on the importance of students' attributions is also abundant (Yeager & Walton, 2011). A range of interventions to help students expect early struggle, to see ability as malleable/incremental rather than fixed/entity or to attribute results to strategy use have found that future expectancies, persistence and performance can be improved by encouraging adaptive attributions (Dweck, 2000; Weiner, 1985; Yeager & Walton, 2011).

Although we are not aware of any direct evidence for this hypothesis, it may be that the teaching skills and behaviours that promote a supportive environment belong in the more advanced end of the teacher development curriculum. It may be possible for competent teachers to be quite effective in promoting learning for most students without really paying much attention to this dimension – that might explain why some of the empirically grounded frameworks, such as the Dynamic Model of Creemers and Kyriakides (2011), do not even include this aspect. Perhaps classroom environment becomes important for determining learning only when other things are well established, or matters significantly for only some students. Overall, it seems unlikely that devoting effort to improving this dimension will be a high-leverage strategy for improving outcomes for most teachers. Nevertheless, we have included it because: (a) there is good evidence that it can have at least a small impact on learning in general classrooms; (b) there may be some contexts or individuals for whom the impact is much larger; and (c) there is good evidence for its impact on wider outcomes, such as student wellbeing and attitudes (Pianta et al., 2012).

03

Maximising opportunity to learn



Great teachers manage the classroom to maximise opportunity to learn

No model of teaching effectiveness could be complete without classroom management: managing the behaviour and activities of a class of students is what teachers do. Yet is it also controversial. Different teachers have very different styles, values and priorities.

Summary of Dimension 3

3.1 Managing time and resources efficiently in the classroom to maximise productivity and minimise wasted time (e.g., starts, transitions); giving clear instructions so students understand what they should be doing; using (and explicitly teaching) routines to make transitions smooth

3.2 Ensuring that rules, expectations and consequences for behaviour are explicit, clear and consistently applied

3.3 Preventing, anticipating and responding to potentially disruptive incidents; reinforcing positive student behaviours; signalling awareness of what is happening in the classroom and responding appropriately

One of the most persistent findings of the 'process-product' classroom observation tradition is the importance of 'opportunity to learn' and 'time on task' (Creemers et al., 2013; Muijs et al., 2014). No teacher will be surprised to learn that being presented with curriculum content in an accessible format and having time to engage with it are found to be necessary for learning to take place. Managing lessons so that time is used productively is a core teaching skill. Specific practices, such as giving clear instructions and establishing routines and rules, support this. Managing student behaviour comes under this heading too: dealing with disruption, but, crucially, preventing it happening in the first place.

Classroom management and efficiency are featured in all the frameworks for teaching quality we have seen, but different frameworks seem to cut this dimension in different ways. We have gone with three strands, though each comprises a collection of techniques, practices and principles: (1) using time efficiently, (2) establishing clear rules, and (3) managing disruption. We also note that how these phenomena are manifested in a particular classroom depends on a lot more than just the skill and behaviours of the teacher: characteristics of the students and the wider school environment/ policies, for example, are both important factors (Bennett, 2017). The same teacher seen teaching a difficult topic to a class containing individuals with persistently challenging behaviours in a school where the rules are unclear or inconsistently applied, might look very different with a fun topic, a class full of biddable students and in a school with strong behaviour support.

Nevertheless, our purpose is not to evaluate teachers but to help them improve. If there are skills that teachers can learn to improve the efficiency, stability and focus of their classrooms, then those skills should be captured in our model so that we can give teachers good feedback about their current status, likely areas and directions for high-leverage improvement effort, and ongoing progress and growth.

Elements of Dimension 3

1. The first element of this dimension relates to the efficient use of time and resources. Great teachers plan activities and resources so that everything works smoothly. Settling down time at the start of a lesson or after a transition is minimised – students get started on meaningful work straight away and work right up to the end of the lesson. Part of this is about giving students clear and simple instructions so they know exactly what they should be doing. Routines can also be an element of great teaching – explicitly teaching students a pattern of behaviour that will be used regularly.

2. The second component is about the consistent and fair application of rules. Rules and expectations should be clearly understood and accepted by all students. Violations should be rare, but when they do happen are treated fairly and appropriately, and as consistently as possible, so that students know that predictable consequences will follow.

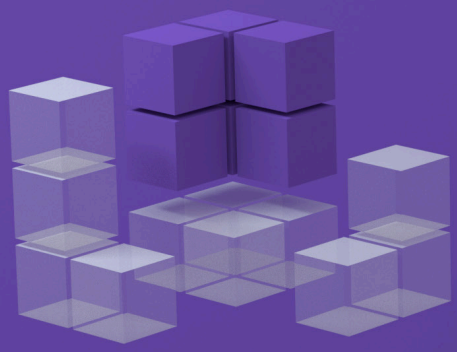
3. The third element concerns preventing and responding to disruption. One of the features of great teaching is that disruption is not seen, but this is often because the teacher has successfully anticipated and prevented it happening. Kern and Clemens (2007) review research on 'antecedent strategies' – whole-class and individually-targeted strategies that teachers can use to "establish a classroom environment that is positive, orderly, predictable and motivating" as a way of preventing disruption and managing student behaviour. The term 'withitness' was coined by Kounin (1977) to describe a teacher's awareness of what is happening in the classroom, even when their attention appears to be elsewhere. Great teachers do not actually have eyes in the back of their head, but their students may think they do. A key part of this skill is that the teacher signals their awareness, perhaps with just a look or movement, so students feel they are under surveillance. Great teachers also use praise and positive reinforcement to support desired behaviour (Calderella et al., 2020). When disruption or disorder does occur, teachers respond firmly and appropriately to minimise the effect on learning. Great teachers draw on targeted approaches that are tailored to the individual needs of students with a history of challenging behaviour.

Evidence for Dimension 3

A large body of evidence supports the use of these strategies to promote learning, so much so that it hardly needs unpacking (for example: Creemers & Kyriakides, 2011; EEF, 2019; Kern & Clemens, 2007; Moore et al., 2019; Muijs et al., 2014, 2018; Scheerens & Bosker, 1997; van de Grift et al., 2017). Praetorius et al. (2018) present evidence of correlations between measures of attainment and classroom management from 1,000 classrooms across ten studies, with a median correlation of 0.18 – the largest of their three dimensions.

04

Activating hard thinking



Great teachers present content, activities and interactions that activate their students' thinking

In many ways, Dimension 4 represents the heart of great teaching: getting students to think hard about the material you want them to learn. It may also be the hardest part of the job to learn, partly because it is rare to get reliable feedback about whether it is working: student learning is invisible, slow and non-linear, so how can we tell if it is happening?

Summary of Dimension 4

4.1 Structuring: giving students an appropriate sequence of learning tasks; signalling learning objectives, rationale, overview, key ideas and stages of progress; matching tasks to learners' needs and readiness; scaffolding and supporting to make tasks accessible to all, but gradually removing them so that all students succeed at the required level

4.2 Explaining: presenting and communicating new ideas clearly, with concise, appropriate, engaging explanations; connecting new ideas to what has previously been learnt (and re-activating/checking prior knowledge); using examples (and non-examples) appropriately to help learners understand and build connections; modelling/demonstrating new skills or procedures with appropriate scaffolding and challenge; using worked/part-worked examples

4.3 Questioning: using questions and dialogue to promote elaboration and connected, flexible thinking among learners (e.g., 'Why?', 'Compare', etc.); using questions to elicit student thinking; getting responses from all students; using high-quality assessment to evidence learning; interpreting, communicating and responding to assessment evidence appropriately

4.4 Interacting: responding appropriately to feedback from students about their thinking/knowledge/understanding; giving students actionable feedback to guide their learning

4.5 Embedding: giving students tasks that embed and reinforce learning; requiring them to practise until learning is fluent and secure; ensuring that once-learnt material is reviewed/revisited to prevent forgetting

4.6 Activating: helping students to plan, regulate and monitor their own learning; progressing appropriately from structured to more independent learning as students develop knowledge and expertise

Elements of Dimension 4

Partly because this fourth dimension is so complex, there seems to be a wide range of different ways to present it in different existing frameworks. We have split it into six elements here, though the total weight of content in this dimension means they are each quite broad and inevitably overlapping. It seems likely that when we start to develop instruments to give teachers feedback about their development, some further splitting may be required. Our six elements are: structuring, explaining, questioning, interacting, embedding and activating.

1

Structuring refers to the choice, matching and sequencing of learning tasks and signalling how they contribute to learning goals. Great teachers share learning aims with their students in ways that help students to understand what success looks like. This does not mean simply writing out lesson objectives or (worse still) getting students to copy them down. Abstract statements of learning aims may be useful but are certainly not enough. To specify learning aims properly, teachers also need to have examples of the kinds of problems, tasks and questions learners will be able to do, as well as examples of work that demonstrates them, with a clear story about how and why each piece of work meets each aim. Great teachers also help students to understand why a particular activity is taking place and how current learning fits into a wider structure. They draw attention to key ideas and signal transitions between activities that focus on different parts of the journey.

A component of structuring is the selection of learning tasks. Tasks must present an appropriate level of difficulty for each student: hard enough to move them forward, but not so hard that they cannot cope, given the existing knowledge and resources they can draw on. Tasks must also promote deep rather than just surface-level thinking (Hattie, 2012), focusing on abstraction, generalisation and the connectedness and flexibility of ideas rather than just reproduction of facts or procedures. In planning a curriculum, tasks must be sequenced so that prerequisite knowledge and skills are accessible and fluent when they are needed. Great teachers build in opportunities for review to check this is the case – and adapt their plans if not.

Great teachers also recognise that complex tasks often require scaffolding: beginning with a simplified or limited version of the task to make it manageable. This often requires some differentiation, as different learners may begin with different levels of readiness and different capacity for learning new material. A knowledge of individual students' needs, including SEND, comes into play here. However, one of the defining characteristics of great teachers is that they require all students to achieve success (Hattie, 2012). Scaffolding provides a gentler entry, but the destination remains the same. Lower-attainers may take longer and need more help, but the job of teachers is to 'disrupt the bell curve', not just to preserve it (Wiliam, 2018). The crucial thing about scaffolding is that you take it away as ideas and procedures become secure and fluent: by the end, those complex tasks are accessible to all.

2

The second element of Dimension 4 is **explaining**. All teachers present new content and ideas to students, but the best presentations have concise, appropriate, engaging explanations that are just right for the students: neither too short nor too long; neither too complex nor too simple. Evidence from both **cognitive load theory** (CLT, Sweller et al., 1998, 2019) and direct instruction (Adams & Engelmann, 1996; Stockard et al., 2018) supports the importance of good explanations. In presenting material, teachers should pay attention to the 'cognitive load' it presents to their students: limiting the number

Cognitive load theory:
Since Sweller first proposed cognitive load theory in the 1980s, it has evolved. A key thread that runs through this research is that humans' capacity for processing information is limited. Working memory can only handle so much at a given moment; this is dependent on the type and complexity of the information. Understanding CLT can be helpful in becoming better at explaining, but there is more to it.

and complexity of new elements; breaking complex ideas or procedures into smaller steps; helping students to assimilate concepts into – and extend – existing schemas; minimising extraneous, irrelevant or distracting input, from either content or environment. Presentations should be planned, crafted and refined, using the collective expertise of experienced teachers and the wisdom of trial and error, to make them as effective as possible.

Part of the skill of explaining is connecting new ideas to prior knowledge. Great teachers know that durable and flexible knowledge depends on connecting ideas together, creating and modifying schemas. A schema is a cognitive structure that enables information to be organised and stored in long-term memory. Schemas are very powerful for learning because they allow individual bits of knowledge to be 'chunked' together into an overarching principle or concept, or for a series of procedures to be combined into a single 'script', and hence processed as a single element. A simple example would be a beginning reader's schema for the letter 'a', which allows them to recognise that a whole range of different shapes (e.g., a, 𝖺, a, 𝐚, 𝖆, A, 𝘈) are actually equivalent in terms of their meaning. The steps in a procedure, such as column subtraction, or conjugating regular -er verbs in the present tense in French, can also be stored as a schema, allowing the whole process to be treated as a single, automated element that can be drawn on in solving a more complex problem. Prior knowledge is structured in schemas and the process of acquiring new knowledge consists of accommodating it into existing or modified schemas and making connections between them (CESE, 2017; Sweller, 1994). Hence, learning depends on the connections that learners make between new ideas and what they already know. Great teachers activate that prior knowledge, reinforce it and connect new ideas to it.

A key insight here is that long-term memory is not just a storage facility, analogous to an encyclopaedia or information searchable on the internet; nor is it limited to routine facts. Instead, the structure and connections among elements of memorised knowledge are precisely what enable it to be used in solving problems or performing complex tasks: if it is not structured and accessible in memory, it cannot be used. Conversely, if a student has a good store of well-structured knowledge, and fluent, automated skills, absorbing new ideas and procedures is much easier. In the same way that gardeners prepare the soil before sowing seeds, great teachers prepare their students for new knowledge by ensuring their existing schemas are well-connected, fluent and accessible. This is one of the reasons why presenting great explanations is not just a generic skill, like being a good communicator: it depends on a detailed knowledge of the content and ideas being explained and how they are learnt.

One of the ways teachers explain new ideas is with the use of examples. Examples can make the abstract concrete and support conceptual understanding if used appropriately (Booth et al., 2017; Braithwaite & Goldstone, 2015). Examples supply content to the theory-building and schema-developing processes that are necessary for new knowledge to be connected, classified and stored. Also necessary for these processes are non-examples and borderline cases: the exceptions and hard cases that define the boundaries of a rule or definition. For learners to construct strong schemas, they need to understand the limits between what does and does not count as an example.

There is also a good deal of evidence that the use of worked examples can be helpful in introducing new ideas (Booth et al., 2017; Sweller et al., 2019). Particularly effective are 'completion problems' where students are given partial solutions and required to complete them. These can help students to focus on the examples but also manage the difficulty level while retaining authentic tasks.

3

Our third element is **questioning**. Pretty much every model of teaching includes this in some form. For example, Rosenshine enjoins us to 'ask a large number of questions and check the responses of all students' (2010, p. 12). But questioning is already one of the commonest things teachers do, and the key to quality is not the number of questions but the type and how they are used. For Hattie (2012) it is about the balance between deep and surface-level thinking that teachers promote. When Smith et al. (2008) searched for the strongest differentiators between 'expert' and 'experienced' teachers they found a focus on promoting deep learning to be one of five distinguishing characteristics (along with: presenting content effectively; creating a learning climate; monitoring and giving feedback; believing that all students can succeed). Hattie (2012) defines this deeper understanding as 'more integrated, more coherent and at a higher level of abstraction'. The key point is that just asking a lot of questions is not a marker of quality; it's about the types of questions, the time allowed for, and depth of, student thinking they provoke or elicit, and how teachers interact with the responses.

This raises an important distinction between different reasons teachers do questioning. Understanding and promoting great teaching requires us to attend to teachers' purposes as well as their practices: not just what they do, but why they do it; what problems they are trying to solve (Kennedy, 2016). Teachers use questioning for two main – and quite distinct – purposes: to promote students' thinking, and to assess it.

In the former purpose, questioning is a tool to promote deep and connected thinking. Great teachers use questioning as part of a dialogue in which students are engaged and stretched. They prompt students to give explanations and justifications for their answers, or just to improve an initial response, to describe their thinking processes, to elaborate on their answers, exploring implications, 'what-if's and connections with other ideas and

knowledge (Dunlosky et al., 2013; Praetorius et al., 2018). Although we have used the word 'questioning' here, the range of activities teachers use to promote oracy and dialogue are much wider. They may also encourage students to ask their own questions. Shimamura (2018) encourages learners to apply the 'three Cs' (categorise, compare and contrast) and 'elaborative-interrogation' (asking, and answering, 'why' and 'how' questions) to help them learn new ideas. Great questioning promotes deep student thinking, helping them to connect and elaborate ideas.

In questioning designed for the latter purpose, the focus is on eliciting and checking student thinking, knowledge and understanding: in other words, assessment. Asking questions, or providing prompts, that provide clear insight into whether students have grasped the required knowledge and understanding is hard; it is in the nature of assessment (and indeed all human communication) that student responses are always equivocal, and interpretations should be probabilistic rather than certain. Questioning that is interactive may go some way to overcome this if follow-ups and prompts are used skilfully to clarify. Great teachers also have strategies for checking the responses of all students. Asking meaningful and appropriate questions that target essential learning, collecting and interpreting a response from every student, and responding to the results, all in real time in the flow of a lesson, is hard to do well, but great teachers do it and it is probably a skill that can be learnt.

Whether questions are asked interactively or as part of a fixed assessment process, starting with great questions that provide maximum information is key. When used for the purpose of assessment, questions should be seen as tools to elicit insights into students' thinking. Questions provide information if they discriminate between those who know and those who don't yet. Whether an assessment is a single question or a formal examination, great teachers understand the amount of information it provides, how much weight it carries and what inferences and decisions it can support. They understand that what has been learnt is not the same as what has been taught (Nuthall, 2007) and that assessment is the only tool we have to make the former visible, albeit 'through a glass, darkly'. Crucially, they plan and adapt their teaching to respond to what assessment tells them.

4

This responsiveness is at the heart of our fourth element, **interacting**. The quality of learning interactions between teachers and students is central to the learning process. Interactions may be seen as a form of feedback, and again there are two distinct purposes here: feedback to teachers that informs their decisions, and feedback to students that helps them learn.

The former purpose, feedback to inform teacher decisions, overlaps considerably with the previous element. Information from questioning and assessment is the basis of this feedback. But it is how the teacher responds to the feedback that matters. First of all, teachers have to understand and interpret the assessment result appropriately. They may need to check or

verify that their interpretations are correct. They also need to appraise the context accurately, being sensitive to the needs, history and dispositions of the student(s) involved. Then they need to identify and decide among a set of options for action. Each will have trade-offs between, for example, time, effort and reward. If some students need more time and help with a topic while others are ready to move on, for example, this may be a hard choice. Finally, they need to implement the chosen option effectively to achieve the desired learning.

For the latter purpose, feedback goes the other way: to the student. Although we know that feedback can enhance learning powerfully (Hattie & Timperley, 2007), we also know that the mediating effects of different combinations of kinds of feedback, learner and task characteristics and different ways of giving feedback are extremely complex. There is no simple recipe for giving powerful feedback. Feedback can help by clarifying or emphasising goals or success criteria ('Where am I going?', Hattie & Timperley, 2007), thus directing students' attention to productive goals. It may draw attention to a gap between actual and desired levels of performance ('How am I going?'), which, again, may be positive if goals are challenging, accepted and accompanied by feelings of self-efficacy (Locke & Latham, 2002). It may cue attributions for success or failure to reasons the student can control, such as effort or strategy choice (Dweck, 2000). Or it may indicate productive next steps ('Where to next?', Hattie & Timperley, 2007). This last mechanism may be the hardest to predict and deliver, precisely because it is a complex interaction between what the learner knows already, what they need to know and their readiness to do what is required to bridge the gap. It also requires an expert judgement about the kinds of actionable next steps that are most likely to deliver the most learning, given all these variables. Great teachers have enough knowledge and experience of similar situations to develop sound intuition about what is likely to work best (Hogarth, 2001), but such intuition is hard to capture in simple rules.

5

The fifth element is **embedding**, getting the learning to stick. The importance of embedding learning rests on the insight from cognitive load theory that memory is not just a storage facility for facts that could just as easily be looked up: the schemas that we use to organise knowledge in memory are the very things we use to think with and to connect new learning to (Sweller, 1994).

There are numerous ways great teachers embed learning. One is by ensuring that students practise any procedures that are regularly required to be fluent and accurate. A large body of psychological research shows that 'overlearning' (continuing to practise after performance has reached a specified standard) can be important for producing learning that is durable and flexible (Soderstrom & Bjork, 2015). Knowledge or schemas that are required for future learning must be secure and readily retrievable. Forgetting is normal but can be slowed or prevented by periodic revisiting and review.

Great teachers ensure that students practise until learning is fluent, automatic and secure.

An important point to note here is that student practice generally needs to be monitored and guided initially (Rosenshine, 2010). In new learning, there is typically a transition: practice begins as helping to learn the ideas, developing connections and understanding, and building schemas; then follows consolidation, gaining confidence and fluency, in which scaffolds and other supports are removed, as is the need for teacher guidance and monitoring; finally comes embedding, where practice becomes independent, fluent, accurate and automatic. Great teachers understand and plan for this transition, monitoring and supporting each student's passage through it and ensuring there is adequate time for each stage.

Practice is particularly effective if it is distributed or 'spaced' over time, with deliberate gaps between for forgetting. Distributing practice like this makes learning feel harder and reduces performance during actual practice, even though it is more effective in the long term – what Bjork and Bjork (2011) have called a 'desirable difficulty'. Great teachers provide opportunities for students to practise procedures and recall of information that must be learnt until it is fluent, and to repeatedly revisit and re-practise after allowing time to forget.

Another approach to embedding is to exploit the 'testing effect', requiring learners to generate answers or recall information from memory in a (low-stakes) test-like process. Again, a vast body of research shows that this is the single most effective way to increase long-term retrieval strength: the ability to recall information or procedures after a delay (Adesope et al., 2017). Moreover, the benefits of testing are not limited to simple recall; the process of having to search for and generate answers also strengthens the connections with, and retrievability of, related information (Delaney et al., 2010). As with all learning, students get better at what they are required to do, so it is important to require them to answer questions that go beyond simple recall and surface-level thinking. Great teachers use the testing effect to delay forgetting with questions that require deep and connected thinking. And of course, testing and spacing can be combined by making time to revisit previously learnt, but about to be forgotten, material after a suitable delay.

There are also other practices that, if done well, can help to ensure learning is durably and flexibly embedded. These include interleaving, varying the conditions of practice, elaboration, and self-explanation (Bjork & Bjork, 2011; Dunlosky et al., 2013; Weinstein et al., 2018). Great teachers understand the principles behind these effects and the contexts in which they are likely to be useful, have a range of strategies for deploying them in practice, and incorporate appropriate and effective use into their teaching.

6

The sixth and final element of Dimension 4 is **activating**: helping students to become independent by planning, regulating and monitoring their own learning. Activating, and in particular promoting, student **metacognition**, is a feature of many of the research-based frameworks (e.g., Ko et al., 2013; Praetorius et al., 2018; van de Grift et al., 2017).

Metacognition:
Although it has a simple literal meaning of "thinking about thinking," metacognition has developed into a broad umbrella term for a number of related cognitive processes. Different frameworks have chosen to focus on different aspects or definitions of this concept. Ultimately, the associated strategies share the aim of helping learners plan, monitor, and evaluate their learning.

When teachers introduce new ideas, it is appropriate to be directive: presenting structured content explicitly, directly teaching what needs to be understood. However, for most educators, the larger aim is to wean students off this dependency on the teacher, encouraging them to become independent, self-actualised learners. In some accounts, this contrast is presented as a polarised opposition between 'traditional', teacher-led, didactic approaches on the one hand, and, on the other, 'progressive', student-focused, constructivist methods and beliefs. In part at least, this division reflects a misunderstanding of the complexity of teaching: different approaches work best at different times, with different students, according to different learning aims, at different stages in the learning process, etc. One approach doesn't fit all.

Within cognitive load theory, both the 'expertise-reversal effect' and the 'guidance-fading effect' refer to the finding that strategies such as presenting limited, structured content and worked examples, which work best for 'novices' (i.e., students who do not yet have the knowledge of the topic or domain encoded in schemas in long-term memory) are no longer the most effective for 'experts', whose chunking and automation of individual elements allow them to tackle and learn more from solving whole problems (Sweller et al., 2019). Using problem-solving as a teaching strategy is overwhelming and inefficient for learners who do not have the required background knowledge, but becomes optimal and necessary when they do.

Interventions to promote the use of metacognitive strategies are among those with the largest effects on attainment, and strategies to help students plan, monitor and evaluate should be explicitly taught and supported (EEF, 2018). Students of all ages should be explicitly taught strategies to plan, monitor and evaluate their learning, ideally in the context of the specific content they are learning. Great teachers also draw attention to their own planning and self-regulation when they model the process of completing complex tasks, and similarly encourage students to 'self-explain' their thinking.

Evidence for Dimension 4

Dimension 4 of our model is derived from the 'cognitive activation' dimension of Praetorius et al. (2018), but it features in every other framework too.

A significant challenge with this dimension is that most of the teacher behaviours that have been found to be effective for activating students' thinking are quite complex. There isn't a simple recipe for developing students' metacognition, for example, or for giving students actionable and appropriate feedback. The very same teacher action could be good in one context, with students of a particular age, history and level of knowledge/skills, in relation to a particular piece of work, and bad in another. Some of the instruments for evaluating teaching make a distinction between 'low-inference' indicators, where the judgement or rating can be made quite easily and 'high-inference', where a more complex specification and a considerable amount of training for raters is required to get consistency. For many of the really powerful elements of cognitive activation, capturing a valid indicator is at the 'high-inference' end of this continuum. It may also be that even expert, trained observers simply cannot perceive enough of the complexity and subtlety of the classroom context to make valid judgements about whether a particular practice is 'good'. This certainly feels like an area where verbal descriptions of practice are inadequate, or at least only a starting point. Understanding what each element means and what really excellent practice looks like could be seen as a life's work.

Another complexity is prioritising among all these elements. Not all of these are important for every teacher to work to improve. It may be, for example, that some parts of the previous three dimensions are prerequisites for this one: if you don't have the content knowledge, or basic classroom management, then those should come first. Some elements of cognitive activation may be a career-long project: even an experienced, expert teacher may find value in improving these aspects of their practice. Some may be best bets for quite a large group of teachers. Wiliam (2018) argues, for example, that a small number of strategies within this dimension, grouped as comprising formative assessment, are likely to offer the highest leverage for most teachers. Rosenshine's (2010) ten principles of instruction may also be seen as high-leverage skills within this dimension.

We think the jury is still out on this question of priorities: existing evidence and theory cannot give an individual teacher a clear-enough steer about which element they should prioritise, or even whether they should try to work on more than one. In the subsequent stages of this project we hope to collect data from teachers working in different ways to improve their practice so that we can learn how to match different kinds of advice, guidance and support to the individual needs of a teacher in order to have the biggest positive impact on student learning.

Structuring is an explicit focus of many of the existing frameworks. For example, it is one of the eight dimensions of the Dynamic Model (Creemers and Kyriakides, 2011), as is 'orientation', which involves clarifying and sharing objectives, and is merged here under the heading of structuring. Careful curriculum sequencing is emphasised in a number of well-validated models of teaching, including mastery learning and direct instruction (Creemers et al., 2013). Selection of appropriate learning tasks and matching their difficulty to students' existing knowledge and readiness, including scaffolding for difficult tasks, features in many models and reviews of effective instruction (e.g., Ko et al., 2013; Muijs et al., 2018; Praetorius et al., 2018; Rosenshine & Stevens, 1986; van de Grift et al., 2017). The need for a balance between foundational knowledge and higher-level extension into 'deep thinking' for all learners is also widely supported (e.g., Hattie, 2012; Pianta et al., 2012; Praetorius et al., 2018).

Support for the importance of explaining draws on evidence from both cognitive load theory (Sweller et al., 1998, 2019) and direct instruction (Adams & Engelmann, 1996; Stockard et al., 2018), as does the use of examples, non-examples, worked examples and completion problems (Booth et al., 2017; Braithwaite & Goldstone, 2015; Sweller et al., 2019). The importance of clear presentation of ideas is an explicit focus of both the ISTOF and ICALT frameworks (Muijs et al., 2018; van de Grift et al., 2017).

Teachers' use of questioning is also widely featured in the evidence-based frameworks (e.g., Creemers & Kyriakides, 2011; Muijs et al., 2018; Rosenshine & Stevens, 1986; van de Grift et al., 2017). Most of these emphasise the importance of the types of questions asked and how teachers respond to them, as do Hattie (2012), Smith et al. (2008) and Ko et al. (2013). The use of elaborative interrogation is judged to have 'moderate utility' by Dunlosky et al. (2013). Questioning as part of formative assessment has a strong evidence base (e.g., Wiliam, 2010).

Interacting denotes the quality of learning interactions between teachers and students, including feedback in both directions. Evidence for the importance of feedback in learning is abundant (e.g., Hattie & Timperley, 2007; Kluger & DeNisi, 1996), especially if combined with goal-setting (Locke and Latham, 2002). Evidence about the role of feedback in prompting adaptive attributions has been cited above under Dimension 2, Element 4.

Embedding learning through practice and retrieval features in some frameworks (e.g., Creemers & Kyriakides, 2011; Rosenshine, 2010) but is noticeably absent from others. These practices draw both theoretical and empirical support from cognitive science, including studies in authentic school classrooms (e.g., Adesope et al, 2017; Delaney et al., 2010; Dunlosky et al., 2013; Weinstein et al., 2018).

Activating – supporting students as self-activated learners – includes strategies that encourage independence, planning, regulating and monitoring. These teacher behaviours are explicitly mentioned in many of the research-based frameworks (e.g., Ko et al., 2013; Praetorius et al., 2018; van de Grift et al., 2017). Extensive evidence from intervention studies supports explicit teaching of metacognitive strategies (e.g., Donker et al., 2014; Hacker et al., 2009).

What next?

We hope that our model can help you make more evidence-informed decisions about how to spend a limited amount of time for professional development to get the biggest return in improving student learning.

Having read through the model, you might be thinking how useful it would be to have examples of these elements, to help anchor and orientate practice in different phases and subjects. We agree, and that's where you come in!

Your profession needs you!

You, like thousands of others, will read this review through a lens of your individual context, phase or subject. It would be simply impossible for us to create accessible examples for everybody and to do them all justice.

So, we welcome you to join the Great Teaching community.

We ask that you share your examples of these elements of Great Teaching, to tell us what they look like in your phase and subject. We ask that you discuss them with other education professionals, to begin reflecting on and improving your practice. Through your insight, you will help us shape the next steps of the Great Teaching Toolkit.

Head to **www.greatteaching.com** to start sharing and get inspired.

Dig deeper into the evidence

You have been presented with a model for great teaching, and a way to get involved with it, but you may be wondering how we got to these conclusions. Where is the evidence behind it?

In the remainder of this report you can go into more depth in each of the four dimensions, learn all about our review methodology and find an overview of all the studies we reviewed.

Appendix 1: Review methodology

This is a review of existing evidence about what teacher competencies (i.e., teacher behaviours, skills, knowledge, beliefs or other abilities) are 'best bets' to be worth learning in order to improve the impact of teaching. As well as these teacher competencies, we are also interested in evidence about environmental proxies that may provide a valid and timely indicator of the quality of student learning taking place in a classroom. For example, if research showed that relationships of trust and respect between students and teachers were predictive of more learning, and that feeding back an indicator of the quality of those relationships could help to improve them, then we might want to include this in our model, even though it is not directly capturing a specific teacher behaviour.

Research questions

1. What teacher competencies (i.e., teacher behaviours, skills, knowledge, attitudes, beliefs or other abilities) are 'best bets' for a teacher to try to learn in order to improve the impact of their teaching?
    a. How have these competencies been captured in existing studies? (Are there adequately reliable and valid measures?)
    b. Are they predictive of student learning? What kinds of relationships have been sought/found (e.g., linear, non-linear or threshold effects)?
    c. What do we know about the dependencies among different competencies? Do they interact or depend on each other?
    d. Is there evidence these competencies can be learnt?
    e. Is there evidence that deliberate attempts to improve the competency lead to greater student learning?

2. What kinds of specific, real-time measures of the quality of a classroom environment may be useful, immediate proxies for the student learning that is occurring?
    a. What evidence supports their use as valid indicators of classroom quality?
    b. Is there evidence that feedback to teachers based on these indicators can help them to improve?

Overview of approach

Systematic review:
In formal terms, a systematic review uses a very explicit and precise procedures to identify, select, appraise, and synthesise existing research – sometimes encompasing hundreds or thousands of studies. This section explains how we took a reasoned, practical approach; while valid, it would not fall into the technical category of a systematic review.

The literature that could potentially be relevant to both these questions is so big and diverse that to attempt a comprehensive, **systematic review** would be a colossal task. We certainly did not have the time and resource for such a project, and it is debatable whether this would be time well spent.

Nevertheless, any review of this kind must address two kinds of threats:

- Comprehensiveness: How do we know we have included everything that is relevant? Have we missed or excluded things that should have been considered or included?
- Bias: Might we have emphasised or favoured perspectives or studies that offer a narrow or particular view, perhaps at the expense of other viewpoints?

Our approach was to conduct a rapid 'umbrella' review (i.e., a review of existing reviews), though in many cases we also reviewed the original studies directly, and our searches generated valuable individual studies as well as reviews. From these studies and reviews we extracted a list of the different teacher competencies that have been cited as related to student learning and the environmental proxies that have been claimed as indicators of classroom quality. For each of these elements we evaluated the quality and relevance of evidence supporting its inclusion in a model of 'What is worth learning for teachers?'.

Identifying relevant studies

We used two main approaches to identifying studies: using known reviews and additional systematic search.

Existing known reviews provide a good place to start in a rapid evidence synthesis. Once we had a list of key reviews, we were able to use backwards (studies they cite) and forwards (later studies that cite them) citation search and related articles search (i.e., studies whose citations overlap). Our starting list included both research reviews and existing frameworks.

1. Reviews:
    a. What makes great teaching? Review of the underpinning research (Coe et al., 2014)
    b. Principles of Instruction (Rosenshine, 2010)
    c. Improving Quality in Education: Dynamic Approaches (Creemers & Kyriakides, 2011)
    d. Effective Teaching: A review of research and evidence (Ko et al., 2013)
    e. State of the art – teacher effectiveness and professional learning (Muijs et al., 2014)
    f. Teacher quality and student achievement (Darling-Hammond, 2000)
    g. Improving students' learning with effective learning techniques (Dunlosky et al., 2013)
    h. Visible Learning for Teachers (Hattie, 2012)

2. Frameworks:
    a. Early Career Framework for England (DfE, 2019)
    b. Enhancing Professional Practice: A Framework for Teaching (Danielson, 2007)
    c. CLASS (Pianta et al., 2012)[3]
    d. ISTOF (Muijs et al., 2018)
    e. ICALT (van de Grift et al., 2017)

In addition, we conducted keyword/topic searches of Web of Science, ERIC and Google Scholar. For Web of Science and ERIC the following string generated 18 and 53 hits, respectively:

("teaching effectiveness" OR "teaching quality" OR "teacher impact" OR "teacher effectiveness" OR "teacher quality" OR "teacher skill" OR "teacher characteristics" OR "pedagogical practice") AND ("learning" OR "attainment" OR "student outcomes") AND ("impact" OR "effect" OR "effects") AND ("systematic review" OR "meta analysis" OR "meta-analysis")

3 See also: **https://curry.virginia.edu/classroom-assessment-scoring-system**

Search strings in Google Scholar are limited to 256 characters and generate thousands of hits, so we used the following string and screened the top 100:

("teaching effectiveness" OR "teaching quality" OR "teacher impact" OR "teacher quality" OR "teacher characteristics") AND ("learning" OR "attainment" OR "student outcomes") AND ("impact" OR "effect") AND ("systematic review" OR "meta analysis")

Results were screened on title and abstract and then obtained and reviewed if they seemed relevant to the research questions above (and had not already been captured from the reviews/frameworks). This was more of an ad-hoc than a systematic process, but it allowed us to check that there were no significant omissions from our evidence base derived from known reviews.

Extracting information

For each claim in each review or study identified from the search process, we attempted to record:

- Type of design/evidence: theoretical, correlational, interventional, experimental
- Types of student outcome captured (and the quality of measures used)
- Types of teacher competency captured
- Types of environmental indicator captured
- Strength of the relationship found (either conditional or unconditional – specify which and conditioned on what)
- Context of the study: location, date, student age range
- Quality of the study and strength/relevance of the claims

In practice, many of these details were not readily available and the data extraction process was less systematic and thorough than we might have achieved with more time and resource. Nevertheless, we believe we were able to achieve sufficient saturation of findings and a good compromise between comprehensive and manageable.

Appendix 2: Overview of studies reviewed

Rosenshine (2010): Principles of Instruction

Rosenshine's (2010) "Principles of Instruction" seems largely to have a similar audience to that of the Great Teaching Toolkit. It focuses on "aspects of language learning and instruction that are universal" and proposes adapting the suggestions to local conditions. The ten principles are derived from three sources:

- Research in cognitive science—how the human brain acquires and uses information, as well as the limits of working memory
- Observation of master teachers—those whose classrooms make the highest gains on achievement tests
- Findings of studies that taught learning strategies

A key consideration for Rosenshine's work is its research base. In the 2010 publication, two "suggested readings" are proffered to further illustrate these claims. However, these pieces of research are not themselves reviews, but more often small-scale, limited interventions or correlational studies. This is not to say that the principles do not come from a large body of literature that supports these practices. If these do exist, however, they are neither directly cited nor signposted. Furthermore, the observational nature of some of the argumentation (e.g., "I once observed a class") potentially belies a systematic, evidence-based argument.

Ultimately, given these reservations, Rosenshine's list reads more as ten specific practices that can be observed in good teachers, rather than broad practices with strong evidence bases. The list comes across as theoretical; it appears to be Rosenshine's (perhaps well-informed) musings. Without greater detail about the outcome measures captured, it remains difficult to further validate his argument. With a focus almost purely on cognitive science, the list does not address any practices pertaining to classroom management, environment, teacher knowledge, etc.

Rosenshine, while presenting principles that on face value seem plausible, leaves a significant gap for offering a more evidence-based argument. Some of his earlier work may offer a more rigorous or systematic approach to the literature (and indeed his somewhat arbitrary selection of "further reading" hints at a deep familiarity with the corpus); however, he has provided no information how earlier conceptualisations and reviews have progressed into these ten principles.

1. Begin a lesson with a short review of previous learning
2. Present new material in small steps with student practice after each step
3. Ask a large number of questions and check the responses of all students
4. Provide models
5. Guide student practice
6. Check for understanding
7. Obtain a high success rate
8. Provide scaffolds for difficult tasks
9. Require and monitor independent practice
10. Engage students in weekly and monthly review

Muijs et al. (2014): 'State of the Art' review

Muijs et al. (2014) offer a very different sort of review than Rosenshine. Grounded firmly in the "best evidence", they enumerate classroom behaviours that are positively related to student achievement. While the reference list is extensive and filled with recognisable names and studies, the authors do not explain any selection criteria or search methodology to collect these resources; an initial reading suggested a "greatest hits" sort of approach.

The authors highlight six "behaviours" that they argue have the strongest research base. Some of these are focused on cognitive processes (e.g., "Opportunity to learn and time on task" and "Instruction and interaction"), while are some are focused on other aspects of teaching (e.g., "Classroom climate" and "Teacher expectations"). They also highlighted notable meta-analyses that seek to quantify effective teaching strategies.

Because most of the studies discussed in the section on effective behaviours focus on "basic skills in English and mathematics", they also explore significant research into self-regulated learning and non-cognitive outcomes (e.g., wellbeing, self-concept, motivation, etc.).

Additionally, the authors discuss the dynamic model of educational effectiveness. A key feature of the model is that numerous levels have an effect on student achievement. Within the teacher level, they highlight the model's eight factors and associated elements—these are observable instruction behaviours. These elements, as they are presented by Muijs et al., are generally broad approaches (e.g., "Dealing with student responses" and "Promoting the idea of modelling"), with a few more specific behaviours (e.g., "Outlining the content to be covered and signalling transitions between lesson parts" and "Analysing data in order to identify student needs and report the results to students and parents").

The authors also include a section on what these mean for teacher professional development—both its implementation and content. This section seems less relevant to the current work of the Great Teaching Toolkit. The article concludes without a clear direction but with "an invitation to dialogue".

Overall, their work appears to have a strong basis in educational research. They admit that much of the evidence is drawn from research in specific fields and with basic skills, with student achievement as a

typical outcome variable. However, they attempt to address this gap with the second section.

Ultimately, this review has two shortcomings. First, it does not present a single, clear framework that unifies the issues raised. Given the broad approach it had to the behaviours and elements it discussed, it couldn't even be seen as an effective "shopping list" of things to do. Even if they were behaviours or elements that were succinct, the lack of a central organisation leaves the reader unsure of which items to extract. Is it content from the six-widely researched behaviours? The eight factors of the dynamic model? The twenty elements that are embedded within the eight factors?

Second, the article is not accessible to teachers. In a literal sense, it is behind a paywall on the Taylor and Francis website. Furthermore, the register and voice of the writing make it clear that the intended audience is not meant to be teachers. The focus ultimately was for researchers—and the call to action was not for teachers to implement this, but for the research community to engage with the topics at hand.

Darling-Hammond (2000)

Darling-Hammond's work is often referenced in literature on teacher effectiveness. Her research builds on previous research on individual teachers' attributes, and shifts the focus to a state-level, aggregate focus. Previously, there was little evidence of a relationship between a teacher's academic ability and students' outcomes (as measured by test scores). Somewhat stronger evidence existed of some correlation between a teacher's subject knowledge and student outcomes, but this was only true up to a point—after a certain point, being an expert in an academic field does not translate to increased student learning. However, a stronger-still relationship had been found between a knowledge of teaching and learning and student performance.

Given this body of research, Darling-Hammond conducted a large-scale survey of school characteristics in states across the United States. She also collected state-level data on elementary math and reading achievement. Because the research focuses on aggregate data (i.e., schools and states), the methodology does not focus on classroom practices or techniques. Among the notable findings is a negative correlation between students living in poverty, English as Additional Language learners, and minority ethnicity students and outcomes. Additionally, teacher quality, as indicated by holding a teaching certification and subject degree, has a positive correlation with student outcomes. Two decades later, these findings may not surprise readers, but they were influential in major educational policies in the US.

While it offers us little by way of classroom practices of effective teaching, it further evidences the development of teachers. Teachers can improve their practice; in doing so, it is crucial for them to achieve certain thresholds to progress to higher levels of effectiveness.

Baumert et al. (2010)

Baumert et al. (2010) examine the concepts of content knowledge and pedagogical content knowledge (PCK) in secondary mathematics teachers in Germany. While previous research had viewed subject knowledge as a unitary concept (e.g., Hill et al., 2004), Baumert et al. explore, both conceptually and empirically, models of the content knowledge and PCK as separate, related concepts. Their assumption is that content knowledge was a sort of threshold prerequisite for PCK, but could not act as a substitute.

The researchers examined a representative sample of Grade 10 mathematics teachers in both the academic and non-academic track in Germany. As an extension to the PISA study, teachers of mathematics answered questionnaires and tests of their teaching knowledge; the questionnaires covered

their background (including training), motivations and beliefs about teaching, and professional beliefs. The tests of teaching knowledge assessed both content knowledge in mathematics and PCK – the latter through open-ended tests of hypothetical situations. Furthermore, teachers submitted homework, tests and classroom tasks to be assessed. Similarly, their students' achievement was measured through tests.

Teachers trained for the academic track show significantly higher scores for their content knowledge and pedagogical content knowledge. This is stronger for content knowledge (greater than one standard deviation) than for PCK; the authors hypothesise this could be because of higher requirements placed on certification, or the higher demands of teachers, on the academic track. These differences remain throughout the teachers' careers. The research finds that the greatest prediction of teachers' content knowledge and PCK scores is the type of teacher training programme attended.

39% of the variance in classes' mathematics achievement (without controlling for the academic or non-academic track) is attributable to pedagogical content knowledge of the teacher. While there are socio-political considerations from the findings, they offer a clear conclusion based on their empirical evidence: teachers' pedagogical content knowledge explains the greatest component of increased student achievement.

- Components of pedagogical content knowledge:
  - Tasks – teachers' ability to identify multiple solution paths
  - Students – ability to recognise students' misconceptions, difficulties, and solution strategies
  - Instruction – teachers' knowledge of different representation and explanations of standard problems
    - Curricular level (indicator of cognitive activation)
    - Individual learning support (teachers providing adaptive explanations)
    - Effective classroom management

Dunlosky et al. (2013)

Dunlosky et al. (2013) produce an extensive monograph that explores ten common learning practices. The selection of these practices is not meant to be exhaustive, but to cover a few that were widely identified as common, as well as a few that would be "easy to use." To this end, it is not meant to provide a complete framework of effective teaching. Instead, it reviews the strengths and shortcomings of each of these practices.

The ten techniques reviewed by the researchers include: elaborative interrogation; self-explanation; summarization; highlighting; keywork mnemonic; imagery use for text learning; rereading; practice testing; distributed practice; and interleaved practice.

For each of these ten, the authors describe key research that explains the technique and how it can be implemented. A notable strength of the article is how they also discuss the generalisability across different contexts—including student characteristics and learning indicators. They offer these reviews and critiques independently, rather than as a unified proposal to implement the entire set.

As a result, they offer different assessments of these ten techniques. Practice testing and distributed practice are deemed to be highly effective practices; elaborative interrogation, self-explanation, and interleaved practice are deemed moderately effective. The remaining five are classified with a low utility level. The researchers, however, do not argue that these techniques should be completely abandoned.

Instead, due to the limited contexts or criteria wherein they appear to be effective, coupled with a lack of rigorous evidence, the conclusion is that their role is limited.

Care should be taken to view these ten techniques as a guide for teachers. While the reviews (as the article reads more as multiple semi-independent reviews rather than one single one) of the literature are extensive, the ten topics discussed do not cover every classroom practice. Indeed, the purpose of the review was not to cover all classroom practices, but certain learning techniques. This exclusive focus on cognitive and learning sciences is certainly important, but does not claim to cover the complete range of an effective teacher's actions. Furthermore, this article is incredibly lengthy and laden with technical references and extensive citations. While this is certainly a strength of this source as an academic piece, it is not in a format that is accessible to most teachers. To Dunlosky's great credit, parallel versions have been created that communicate key findings in a medium and style that is accessible and useful to teachers.

The reviews effectively offer researchers some techniques that compose effective teaching, but ultimately stop short of suggesting all the practices effective teaching comprises.

Praetorius et al. (2018)

Praetorius et al. (2018) present a framework for teaching quality that has been widely used in German-speaking countries and was originally developed in the context of maths education for the 1995 TIMSS-Video study. A three-dimensional model emerged from factor analysis of these instruments. The main framework consists of the three main dimensions, beneath which are 21 sub-dimensions. These sub-dimensions are derived from a set of classroom observation scales developed in Germany in the 1990s (Clausen, 2002; Gruehn, 2000). For each sub-dimension, Praetorius et al. give up to three example items to illustrate how it has been operationalised.

One feature of this model is that it contains nothing that is subject-specific: "the dimensions are conceptualised as being generic in nature, and thus as being applicable across school subjects" (p. 2).

The Three Basic Dimensions framework is derived from a theoretically-guided view of teaching and learning, as much as by direct empirical evidence. For example, its view of motivation comes from Deci and Ryan's (2008) self-determination theory, focusing on competence, autonomy and relatedness as the requirements for students to be motivated.

The Three Basic Dimensions Framework

The three main dimensions are classroom management, student support and cognitive activation. Their components, as listed by Praetorius et al. (2018) are set out here:

Classroom management

- (Lack of) disruptions and discipline
- (Effective) time use/time on task
- Monitoring/'withitness'
- Clear rules and routines

Student support

- Support of competence experience
  - Differentiation and adaptive support
  - Pace of instruction
  - Constructive approach to errors
  - Factual, constructive feedback/appreciation
- Support of autonomy experience
  - Interestingness and relevance
  - Performance pressure and competition (negative indicator)
  - Individual choice options
- Support of social relatedness experience
  - Teacher → student
  - Student → teacher
  - Student → student

Cognitive activation

- Challenging tasks and questions
- Exploring and activating prior knowledge
- Exploration of the students' ways of thinking/elicitation of student thinking
- Receptive/transmissive understanding of learning of the teacher (negative indicator)
- Discursive and co-constructive learning
- Genetic-socratic teaching
- Supporting metacognition

Supporting evidence

Praetorius et al. cite 39 research reports, based on 21 research studies/projects, in support of the framework. For each sub-dimension in the framework, the number of studies included and reported is shown in Table 1.

Table 1: Number of studies included in each element of the three-dimensional model in Praetorius et al. (2018)

		No. of cited studies that included this element
Classroom management		
	(Lack of) disruptions and discipline	17
	(Effective) time use/time on task	15
	Monitoring/'withitness'	6
	Clear rules and routines	5
Student support		
	Differentiation and adaptive support	12
	Pace of instruction	6
	Constructive approach to errors	11
	Factual, constructive feedback/appreciation	4
	Interestingness and relevance	5
	Performance pressure and competition (negative indicator)	4
	Individual choice options	6
	Teacher → student	14
	Student → teacher	4
	Student → student	6
Cognitive activation		
	Challenging tasks and questions	16
	Exploring and activating prior knowledge	7
	Exploration of the students' ways of thinking/elicitation of student thinking	8
	Receptive/transmissive understanding of learning of the teacher (negative indicator)	2
	Discursive and co-constructive learning	5
	Genetic-socratic teaching	3
	Supporting metacognition	2

We can see that no sub-dimension was included in all the studies and only a handful were in more than half.

Praetorius et al. also provide evidence of the predictive validity of the elements of the framework: the extent to which they predict learning gains and other outcomes. Correlations are reported at the level of the three basic dimensions (i.e., classroom management, student support, cognitive activation). These dimensions have been measured in a number of different modes: from classroom observation, from student surveys, from teacher self-report, from analysis of classroom artefacts. The outcomes used also vary across studies, ranging from before-and-after standardised curriculum assessments to self-reports of student enjoyment. For the before-and-after measures, the time gap between them also varies: from nine lessons, to one year.

If we limit our focus to studies where the outcome is gains on some kind of assessment of attainment, there are 25 level-2 (classroom) regression coefficients, ranging from -0.27 to 0.46, with median coefficients of 0.18, 0.12 and 0.17 for classroom management, student support and cognitive activation, respectively.[4]

Overall, predictive validity is quite low, and mixed. Even when the best measures are combined, collectively these measures of classroom quality do not explain much of the variation in student learning gains. The authors themselves note that "the results regarding the predictive validity of the Three Basic Dimensions are not convincing" (p. 16).

The authors also acknowledge that there may be other important characteristics of effective teaching not captured in the framework. Gaps might include generic and content-specific elements (p. 17).

All in all, this study offers a useful contribution to the Great Teaching Toolkit. It is grounded in strong empirical data and sound theory, and has been robustly tested in both research and practice. Its evidence base is purely correlational, though it does draw on a range of methods (e.g., observation, student survey), and while correlations are not large (0.1 – 0.2), they are consistent with other studies.

The work of Praetorius et al. makes the case for three broad dimensions in the Great Teaching Toolkit, which seems to represent a reasonable consensus across many other studies. However, the exact contribution of each of the sub-dimensions is less clear. They are certainly useful as exemplars and operationalised elements; whether they are all important for teacher development is unknown.

Seidel and Shavelson (2008)

This meta-analysis makes a number of contributions to our knowledge about the characteristics of effective teaching.

First, as a systematic review and meta-analysis of studies between 1995 and 2004, it provides an important summary of the evidence from that period, though it also sets out to problematise the whole endeavour of meta-analysis. Evidence is synthesised from 112 publications, containing 1,357 estimates of the relationship between a range of instructional or classroom components and student outcomes, adjusted for background characteristics.

4 Taken from Table 4 in Praetorius et al., 2018.

Second, it gives a contextualised summary of and comparison with the results from two previous systematic reviews, by Fraser et al. (1987) and Scheerens and Bosker (1997). While the latter study reported an overall correlation of 0.25 between observable teaching quality and outcomes, Seidel and Shavelson, a decade later (and using a similar theoretical model), report an average of 0.02, with no individual component achieving a correlation higher than 0.04. Their main explanation for this difference is not that the strength of relationship has reduced, but that the later review used better controls[5] and weighting of component study estimates – both of which are standard practice in high-quality meta-analysis today.

Third, it presents, as an alternative to the traditional 'process-product' approach, a more theoretically-guided cognitive model of teaching and learning as a conceptual framework for the meta-analysis. This model focuses on the functions, purpose and context of different classroom components. These are identified as a set of contextual factors and a set of teacher practices, taken from a model outlined by Bolhuis (2003). The contextual factors are: (1) knowledge domain (the subject or curriculum being taught); (2) time for learning; (3) organisation of learning (classroom management); and (4) social context (social learning climate). The teacher practices are: (5) goal-setting/orientation (e.g., clarifying goals, teaching in a clear and structured way, activating student pre-knowledge); (6) execution of learning activities (providing opportunities for processing information, such as modelling, problem-solving and inquiry); (7) evaluation (assessment); and (8) regulation and monitoring (giving feedback and promoting meta-cognitive strategies). This framework is used to classify the effects from different studies and more variation is found, particularly when the 'effects' are separated according to the method of measuring the instructional and classroom components (whether by teacher survey, student survey or observation/video analysis) or split by type of design (experimental/quasi-experimental vs correlational).

Seidel and Shavelson's main substantive finding is that "we found the largest teaching effects for domain-specific components of teaching—teaching components most proximal to executive learning processes". However, it is not very clear exactly what kinds of teacher practices were classified under that heading, beyond the examples cited of "activities such as mathematical problem-solving, scientific inquiry, or specific reading and writing strategies". "Organisation of learning" (i.e., classroom management) was also found to have a consistently strong relationship with learning outcomes.

Creemers and Kyriakides (2006; 2011): Dynamic Model

Creemers and Kyriakides (2006; 2011), and their Dynamic Model, come from the tradition of 'Educational Effectiveness Research' (EER, a blending of previously separate traditions of School Effectiveness Research with studies of classroom practices and teacher effectiveness, some of which go back to the 1960s; Creemers & Kyriakides, 2015; Creemers et al., 2013). This research tradition is characterised by the use of assessments of student learning, often limited to assessments of maths and reading in standardised tests, and statistical models (often multi-level regression models) to adjust for a range of covariates, interpreting the unexplained variation in outcomes as the 'effect' of the school or teacher.

The Dynamic Model takes this work forward by recognising greater complexity in a number of ways. First is the insight that understanding the relationships between educational inputs (resources, behaviours,

5 In particular, they did not use correlations between an observed practice and raw student outcomes unless some adjustment for prior covariates was made (e.g., prior attainment or SES).

policies, etc.) and outputs (e.g., learning) requires a range of different types of theory. In particular, if we want to improve as well as understand, this theory is even more important – especially, good theories of teaching and learning. Related to this, the authors note that they draw on a wide range of different perspectives on pedagogy, including elements that would generally be associated with both direct instruction approaches and constructivism (Creemers et al., 2013).

The weaving together of these two strands is a key element of the Dynamic Model. On the one hand, Creemers et al. (2013) characterise the 'competency-based approach' as setting out to list explicit strategies and competences, drawing on the 'process-product' tradition of work by researchers such as Good and Brophy (1984) and Rosenshine (1976), who observed consistent empirical associations between certain observable teacher behaviours and student achievement. This leads to a view of effective teaching as skills-based and discrete skills, such as classroom management, clear and concise exposition of ideas, and using questioning, modelling, reviewing and feedback. Theories such as Carroll's (1963) model of learning and cognitive load theory (Sweller et al., 1998) are often invoked in this tradition and approaches such as mastery learning (Bloom, 1976) or direct instruction (Rosenshine, 1987) can be seen as practical, packaged instantiations of it.

On the other hand, Creemers et al. (2013) contrast this with a more holistic approach to understanding teaching quality. This tradition draws on the ideas of 'reflection in action' derived from Dewey (1933), Stenhouse (1975) and Schon (1983), that stress the need for teachers' own critical reflection on their practice, and the work of developmental psychologists such as Piaget (1972), who emphasise the need for learners (including teachers learning their practice) to actively construct meaning from experience. Becoming a great teacher requires more than just practising isolated techniques: each teacher must understand their own context, reflect on their practice, and, through action research and teacher inquiry, find their own solutions to the problems it presents. Teaching cannot be reduced to a mechanistic, technicist project: it is just too complex to have a single right way that can be simply described and applied universally. Professional development should emphasise critical reflection, inquiry, agency and moral purpose.

While these two approaches are often seen as incompatible paradigms – and even in opposition, dividing educators between 'traditional' and 'progressive' camps – for Creemers et al. (2013), they each offer part of the story. Research evidence is clear that certain practices and techniques are powerful determinants of student learning: teachers should know about these competences and should be supported in learning them. However, a great teacher is more than just a set of isolated competences: teachers must understand the underlying theory and processes, as well as knowing their own context, so that they can make judicious selections and adaptations, and implement these practices faithfully and effectively. Moreover, according to the evidence from the Dynamic Model, the emphasis of teachers' professional learning should reflect the stage of their development: for less effective (typically early-career) teachers, developing the basic skills is the priority; once those skills are mastered, they have something to reflect with and can apply that reflection to the more complex challenges of refining, orchestrating and adapting their practices.

The second insight of the Dynamic Model is that there are a range of outcomes of interest, reflecting overlapping educational aims – for example, cognitive, psychomotor, metacognitive and affective – not just basic skills. Moreover, some of the validation studies have evaluated the equity of educational outcomes as well as their overall levels. Third, the factors influencing effectiveness operate at multiple levels (student, classroom, school and system). Fourth, these factors may also interact, either within a level or across levels. Some characteristics of effectiveness may act more like catalysts or barriers, in

combination with other factors: the effect of each factor may depend on other factors. Fifth, and related to the previous point, some factors may be expected to have non-linear relationships with outcomes. For example, it could be the case that more of a particular factor is associated with better outcomes up to a point, but beyond that the relationship plateaus or even reverses.

The model specifies what these factors are, based on empirical evidence and tested theory. Where other models typically focus on measuring the amount of a factor, the Dynamic Model recognises that the factors differ qualitatively as well as quantitatively. For each factor, as well as its frequency (how much or how often it occurs), the model captures its focus (its function, specificity and purpose), stage (the duration and timing of a policy or practice), quality (whether it is done well, in ways that are aligned with best evidence) and differentiation (the extent to which its implementation is adapted to individual student/classroom/school context and needs).

At the student level, the Dynamic Model includes the following (Creemers & Kyriakides, 2011, p. 29):

- Socio-cultural and economic factors, such as socioeconomic status, ethnicity and gender, their interactions and compositional effects
- Psychological characteristics: aptitude, motivation, expectations, personality and thinking style[6]
- Opportunity to learn: the time made available and focused on learning activities aligned with the intended curriculum
- Time on task: time actually spent engaged in learning, limited by individual student interest, motivation, expectations, attention, self-regulation and quality of focus

Classroom factors relate to the behaviour of the teacher and 'refer to observable instructional behaviours of teachers in the classroom rather than on factors that may explain such behaviours (e.g., teacher beliefs and knowledge and interpersonal competences)'. There are eight factors in the model:

6 The justification for 'thinking style' draws on the work of Sternberg (e.g., 1988) that tries to account for differences in performance and 'intelligence' in terms of intellectual styles of mental self-government.

Table 2: The dynamic model of educational effectiveness (Creemers & Kyriakides, 2011, p. 35)

(1) Orientation	(a) Providing the objectives for which a specific task/lesson/series of lessons takes place
	(b) Challenging students to identify the reason why an activity is taking place in the lesson
(2) Structuring	(a) Beginning with overviews and/or reviews of objectives
	(b) Outlining the content to be covered and signalling transitions between lesson parts
	(c) Drawing attention to and reviewing main ideas
(3) Questioning	(a) Raising different types of questions (i.e., process and product) at appropriate difficulty level
	(b) Giving time for students to respond
	(c) Dealing with student responses
(4) Teaching modelling .	(a) Encouraging students to use problem-solving strategies presented by the teacher or other classmates
	(b) Inviting students to develop strategies
	(c) Promoting the idea of modelling
(5) Application	(a) Using seatwork or small-group tasks in order to provide needed practice and application opportunities
	(b) Using application tasks as starting points for the next step of teaching and learning
(6) The classroom as a learning environment	(a) Establishing on-task behaviour through the interactions they promote (i.e., teacher–student and student–student interactions)
	(b) Dealing with classroom disorder and student competition through establishing rules, persuading students to respect them and using the rules
(7) Management of time	(a) Organising the classroom environment
	(b) Maximising engagement rates
(8) Assessment	(a) Using appropriate techniques to collect data on student knowledge and skills
	(b) Analysing data in order to identify student needs and report the results to students and parents
	(c) Teachers evaluating their own practices

The Dynamic Model has a number of very specific strengths.

First is the fact that the model is derived from a large body of previous evidence about the effectiveness of schools and teachers.

Second, the Dynamic Model has been subjected to a fair level of rigorous testing of its claims and predictions, albeit mainly by the authors themselves. For example, a 2013 meta-analysis by Kyriakides et al. broadly supports this framework; Creemers et al. (2012) also provide a review of evidence from several studies.

Third, the model is accompanied by an extensive suite of instruments to measure its various components. Creemers and Kyriakides (2011) provide details of these instruments that include student questionnaires, classroom observation schedules (high and low inference), and teacher questionnaires.

A fourth strength is that the implications of the Dynamic Model for school and teacher improvement have been developed and tested in practice. In the study reported by Creemers et al. (2013), 130 primary school teachers in Cyprus were randomly allocated to either a development programme based on the Dynamic Model (the 'Dynamic Integrated Approach' – DIA) or an alternative 'holistic approach', based on supporting teachers' critical reflection on their pedagogy. Teachers were assessed at the start as fitting one of five developmental stages, and randomisation was 'blocked' within each stage (i.e., each stage had equal numbers in each treatment arm). A brief description of the characteristics of the five stages is given in Table 3. Both arms had eight PD sessions, approximately once a month over a school year, in which they developed an individual action plan and received support from peers as well as from expert outsiders, who also observed and gave feedback.

In the DIA treatment arm, teachers were grouped according to their developmental stage and given materials and training specific to their stage. In developing their action plans, they were encouraged to focus on stage-appropriate, research-backed approaches and their progress was supported with ongoing provision of targeted reading materials and tasks. By contrast, teachers in the 'holistic approach' arm were free to choose their own goals and activities, supported by peer-group and expert-moderated discussion and critical reflection.

A number of outcomes were recorded. In terms of observational assessments of teaching quality, about a third of the DIA teachers moved up a stage, compared with none of those in the holistic arm. All teachers in the DIA group improved their quality scores at least to some extent and the mean change was equivalent to about 0.6 SD (compared with 0 for the 'holistic' group). Neither group exhibited any change in their attitudes or perceptions. Assessments of student learning showed an effect size of 0.24 in favour of those in the DIA arm, with larger effects (an additional 0.14) for those teachers who progressed a whole stage. A one-year follow-up of both groups found that the change (or lack of change) in teaching quality had been maintained (Creemers et al., 2013, p. 218).

Creemers et al. (2013) interpret these results as showing that "reflection is more effective when teachers' priorities for improvement are taken into account and when they are encouraged to develop action plans which address their professional needs; these were identified through a relevant empirical investigation". They note that some of the teachers in the holistic arm selected areas to work on that were well above their stage, for example, trying to differentiate their instruction when their classroom management skills were weak: "their attempts to incorporate this into their teaching were not successful" (p. 178).

Table 3: Five stages of teaching skills (based on Creemers et al., 2013, pp. 163, 179)

Stage	Teaching skills	Focus for development
1. Basic elements of direct teaching	• Management of time • Use of structuring, application, assessment and questioning • Some attention to teacher-student relations	*Maximising opportunity to learn* • Lesson structuring (sequencing, connecting to previous, key points explained) • Use of application activities/exercises (practising application and implementation of knowledge/procedures, giving individual feedback and asking questions) • Questioning and providing feedback (asking many questions of all students, giving them time to reflect)
2. Putting aspects of quality in direct teaching and touching on active teaching	• More sophisticated use of structuring, application, questioning (i.e., appropriate for stage and purpose in relation to context) • Some attention to student relations	*Matching lesson activities to purpose and context* • Timing of the application tasks (judging when to use and what knowledge, skills, applications and links to previous learning to include) • Quality of the lesson structuring (effective connections with and review of previous learning, highlighting key points)
3. Acquiring quality in active/direct teaching	• More sophisticated attention to teacher-student and student-student relations (appropriate timing and purpose) • More sophisticated use of assessment, orientation, feedback, questioning and structuring (in relation to timing and purpose, but also quality) • Use of teaching modelling	*Developing the classroom learning environment for constructing knowledge* • Orientation to learning aims (ensuring students understand learning goals) • Development of the classroom as a learning environment (regular, high-quality, learning-focused interactions between teacher and students, and between students; students are encouraged to express their views or explore different solutions, but also challenged to justify them)
4. Differentiation of teaching	• Appropriately differentiated and high-quality use of structuring, time management, questioning, application, assessment, teaching modelling and orientation	*Differentiating appropriately* • Differentiation of teaching (teachers adapting their approach to student characteristics, readiness and needs, in their questioning, feedback, assignment of application tasks and follow-up) • Orientation of students to the learning goals (strategies for engaging students in identifying learning objectives: 'why are we doing this?')

5. Achieving quality and differentiation in teaching using different approaches	• Use of differentiated and high-quality practices in relation to teacher-student and student relations • Appropriately differentiated, appropriate and high-quality use of orientation and teaching modelling

In a partial replication of this study, Creemers et al. (2013) compared groups of teachers in a four-arm randomised trial with a 2x2 factorial design. As before, two of the groups used the DIA approach and two the holistic approach, but this time one of each was supported by external experts (as above) while the other was supported by colleagues within their own school, using a programme designed by the researchers. Results were very similar to the previous study, with gains in observed teaching competence and student attainment for the DIA group, but none for the 'holistic' approach. Interestingly, there was no difference between teachers supported directly by external experts and those supported by their colleagues.

A further evaluation of the DIA (Creemers et al., 2013) compared training a group of teachers in their knowledge and use of assessment. Again, the DIA group were assessed (this time on their use of assessment, using a self-report questionnaire), divided into groups based on their stage and given learning materials and support that were targeted at their level. In this study, the comparison was with a randomly equivalent group allocated to be given skills-based training[7] in assessment, but where the training was the same for all, irrespective of their existing level of practice; there was also a third, no-treatment, control arm. Assessments of their use of assessment (from teacher self-report) improved for both the active groups, but not for the controls. Gains for the 'targeted' (DIA) group were about twice the size of those for the generic skills-based training. There were also gains in student achievement: for teachers judged to be at stage 1 (of 4) in their use of assessment, both the active groups saw improved attainment equivalent to an effect size of about 0.1. For teachers at the higher stages, only the DIA group improved (ES=0.17).

The main takeaway from these professional development studies using the Dynamic Model seems to be that we can usefully identify stages of teacher effectiveness. Despite the diversity of different elements of classroom quality, they seem to cluster into levels. The impact of professional development on student learning seems to be greatest when it is targeted at developing practices and skills that take each teacher from their current stage to the next.

Scheerens et al. (2007) meta-analysis

Scheerens et al. (2007) present a review, conceptualisation and meta-analysis of both school-level and classroom-level factors related to student outcomes. They identify 46 factors at the classroom level and categorise them into 15 teaching dimensions. The mean correlation[8] and number of effects are shown in Table 4.

7 Derived from the competency-based approach, described above.

8 Technically, these are Fisher-Z scores. For the range of values here, they are within 5% of the corresponding correlations.

Table 4: Effect sizes (Fisher-Z) of classroom factors from Scheerens et al. (2007)

Classroom factor	Mean Effect	No of effects	
1	**Learning Time**	**.095**	
1.1	time on task	.125	86
1.2	opportunity to learn	.118	32
1.3	homework	.041	51
1.4	mastery learning	.047	4
2	**Classroom organisation**	**.075**	
2.1	classroom management	.088	36
2.2	discipline	.070	20
2.3	control	.018	17
3	**Learning environment**	**.129**	
3.1	classroom climate	.125	107
3.2	no achievement pressure	.151	29
3.3	mastery orientation	-.005	2
3.4	no performance orientation	.120	2
4	**Clear and structured**	**.126**	
4.1	structured/direct teaching	.107	76
4.2	goal-directed/clear	.222	36
4.3	teacher demonstration	.014	17
4.4	teaching basic skills	.073	17
5	**Activating**	**.123**	
5.1	cooperative	.204	49
5.2	situated/discovery	.155	3
5.3	peer tutoring	.218	53
5.4	student work	.059	36
5.5	individual work	-.009	39
5.6	student discussions	.043	8
6	**Learning strategies**	**.213**	
6.2	meta-cognitive	.244	35
6.3	scientific inquiry	.197	32
6.5	organising methods	.000	2
6.7	reading/writing	.210	34
7	**Challenge**	**.130**	
7.1	cognitive activation/understanding orientation	.182	67
7.2	active student engagement	.042	63
7.3	authentic contexts/relevance	.160	47
7.4	language level	.029	7
7.5	representation formats	.385	4
8	**Support**	**.108**	
8.1	quality of interactions/teacher support	.108	73

9	**Feedback**	**.056**	
9.1	feedback/frame of reference/monitoring	.056	106
10	**Evaluation**	**.086**	
10.1	assessments/tests	.086	46
11	**Teacher characteristics**	**.146**	
11.1	high expectations	.124	22
11.2	constructivist beliefs about learning	.354	4
12	**Adaptive Teaching**	**.066**	
12.1	various teaching methods	.124	2
12.2	adaptive teaching	.036	27
12.3	open tasks/choice	.090	4
12.4	student prerequisites	.178	7
13	**Practice**	**-.080**	
13.1	drill/repetition	-.078	17
13.2	application	-.057	19
14	**Material**	**.015**	
14.1	textbooks	.039	6
14.2	media	.012	27
15	**Integrative approaches**	**.089**	
15.1	constructivist	.039	52
15.2	inductive	-.197	5
15.3	concept-oriented	.257	33

Scheerens et al. note some interesting results. The largest individual factors (e.g., representation formats and constructivist beliefs about learning) come from quite small numbers of studies so should be interpreted cautiously. Among those with more replications, teaching meta-cognitive strategies, peer tutoring, cooperative learning, and instruction that is clearly goal-directed have relatively high coefficients (above 0.2). Similarly high coefficients are also found for subject-specific learning strategies, like scientific inquiry and reading and writing. A number of these larger effects are for factors associated with 'constructivist' approaches, compared with negative effects for practice (drill/repetition, application) and for performance pressure (i.e., positive correlations for no achievement pressure and no performance orientation).

A comparison that groups 'constructivist-oriented' approaches (cooperative, situated/discovery, peer tutoring, student work, individual work, student discussions, meta-cognitive, scientific inquiry, organising methods, reading/writing, cognitive activation/understanding orientation, active student engagement, authentic contexts/relevance, constructivist beliefs about learning, constructivist, inductive, concept-oriented) against 'structured/direct/mastery' approaches (mastery learning, **mastery** orientation, structured/direct teaching, goal-directed/clear, teacher demonstration, teaching basic skills, drill/repetition, application) narrowly favours the former.

Danielson (2007): Framework for Teaching

The Framework for Teaching was created by Charlotte Danielson and particularly grew in popularity during the 2010s. It presents four domains, divided into a total of 22 components. They are: planning and preparation, classroom environment, instruction, and professional responsibilities. In each of these domains, teachers can be labelled as unsatisfactory, basic, proficient, or distinguished.

The *planning and preparation* domain refers not just to the design and preparation of teaching, but also the content knowledge that the teaching entails. The components of this domain also include the implementation of assessments, as well as instruction that is aligned to the curriculum. The second domain, *classroom environment*, refers to both the physical space and the classroom as a social space. This then includes appropriate student behaviour as a critical component. *Instruction* refers to 'engaged learning', which covers students actively engaging with materials at a high level. There are also elements of student metacognition in this domain, with students understanding learning goals. The final domain, *professional responsibilities*, include additional responsibilities placed on teachers that contribute to a school's success. These range from record-keeping, communicating with families, and professional development.

The Framework for Teaching includes a range of behaviours and expectations that extend beyond a definition of effective *teaching*; it could be said they aim to describe a good *teacher*. That is, the professional aspects that extend beyond formal learning (e.g., maintaining accurate records) bear as much weight as instructional ones (e.g., engaging students in learning).

The framework is widely accessed by school leaders; in some cases, it is used as the basis for teacher evaluations. While resources for teachers and leaders can be readily obtained online, greater detail can be found in Danielson's book. The domains and components were developed by both 'practice wisdom' and underlying research; this research, however, is not as readily available as the framework itself.

Early Career Framework (2019)

The Department for Education in England consulted widely with researchers, as well as the wider education sector, to develop the Early Career Framework. In support of this, the Education Endowment Foundation acted as an independent reviewer to ensure the framework is robust and evidence-based. Explicitly not intended to be an assessment rubric, the framework is meant to support and "underpin" professional development for early career teachers.

The framework has eight dimensions. Each enumerates a series of statements of what teachers should learn related to the dimension, as well as the behaviours they should exhibit. The framework drew the "learn that..." statements from the best available evidence; the "learn how to..." statements were additionally collected from expert practitioners.

- **High expectations**
  Learn how to...
  - Communicate a belief in the academic potential of all pupils
  - Demonstrate consistently high behaviour expectations
- **How pupils learn**
  Learn how to...
  - Avoid overloading working memory
  - Build on pupils' prior knowledge

    - Increase likelihood of material being retained
- **Subject and curriculum**
  Learn how to...
    - Deliver a carefully sequenced and coherent curriculum
    - Support pupils to build increasingly complex mental models
    - Develop fluency
    - Help pupils apply knowledge and skills to other contexts
    - Develop pupils' literacy
- **Classroom practice**
  Learn how to...
    - Plan effective lessons
    - Make good use of expositions
    - Model effectively
    - Stimulate pupil thinking and check for understanding
- **Adaptive teaching**
  Learn how to...
    - Develop an understanding of different pupil needs
    - Provide opportunity for all pupils to experience success
    - Meet individual needs without creating unnecessary workload
    - Group pupils effectively
- **Assessment**
  Learn how to...
    - Avoid common assessment pitfalls
    - Check prior knowledge and understanding during lessons
    - Provide high-quality feedback
    - Make marking manageable and effective
- **Managing behaviour**
  Learn how to...
    - Develop a positive, predictable and safe environment for pupils
    - Establish effective routines and expectations
    - Build trusting relationships
    - Motivate pupils
- **Professional behaviours**
  Learn how to...
    - Develop as a professional
    - Build effective working relationships
    - Manage workload and wellbeing

The framework includes both classroom- and learning-focused behaviours (with significant focus on learning and cognitive actions), as well as those of a teacher. It is responsive to current trends and needs in education in England, with workload appearing in multiple capacities. Similarly, understanding how pupils learn is the focus of a whole dimension on its own. Naturally, a strength of the framework is the inclusion of extensive research and resources, each linked to a corresponding dimension. It therefore effectively lays out a series of focuses for early career teachers for their own professional development.

References and further reading

Adams, G., & Engelmann, S. (1996). *Research on direct instruction: 25 years beyond DISTAR.* Educational Achievement Systems.

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659–701. https://doi.org/10.3102/0034654316689306

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science, 333*(6045), 1034–1037. https://doi.org/10.1126/science.1207998

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching. *Journal of Teacher Education, 59*(5), 389–407. https://doi.org/10.1177/0022487108324554

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development, 67*(3), 1206–1222. https://doi.org/10.1111/j.1467-8624.1996.tb01791.x

Baumert, J., & Kunter, M. (2013). The COACTIV Model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV Project* (pp. 25–48). Springer. https://doi.org/10.1007/978-1-4614-5149-5_2

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180. https://doi.org/10.3102/0002831209345157

Bennett, T. (2017). *Creating a culture: How school leaders can optimise behaviour*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/602487/Tom_Bennett_Independent_Review_of_Behaviour_in_Schools.pdf

Bjork, E. L., & Bjork, R. A. (2009). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher & J. R. Pomerantz (Eds.), *Psychology and the real world: essays illustrating fundamental contributions to society* (pp. 55–64). Worth Publishers.

Blömeke, S., Busse, A., Kaiser, G., König, J., & Suhl, U. (2016). The relation between content-specific and general teacher knowledge and skills. *Teaching and Teacher Education, 56*, 35–46. https://doi.org/10.1016/j.tate.2016.02.003

Bloom, B. S. (1976). Human characteristics and school learning. In *Human characteristics and school learning*. McGraw-Hill.

Bolhuis, S. (2003). Towards process-oriented teaching for self-directed lifelong learning: A multidimensional perspective. *Learning and Instruction, 13*(3), 327–347. https://doi.org/10.1016/S0959-4752(02)00008-7

Booth, J. L., McGinn, K. M., Barbieri, C., Begolli, K. N., Chang, B., Miller-Cotto, D., Young, L. K., & Davenport, J. L. (2017). Evidence for cognitive science principles that impact learning in mathematics. In D. C. Geary, D. B. Berch, R. J. Ochsendorf, & K. M. B. T.-A. of C. A. S. and H.-O. M. C. Koepke (Eds.), *Acquisition of complex Arithmetic skills and higher-order mathematics concepts* (pp. 297–325). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-805086-6.00013-8

Bowlby, J. (1969). Attachment and loss: Attachment (Vol. 1). In *Attachment* (Vol. 1). Basic Books. https://doi.org/978/0712674713

Braasch, J. L. G., Goldman, S. R., & Wiley, J. (2013). The influences of text and reader characteristics on learning from refutations in science texts. *Journal of Educational Psychology, 105*(3), 561–578. https://doi.org/10.1037/a0032627

Braithwaite, D. W., & Goldstone, R. L. (2015). Effects of variation and prior knowledge on sbstract concept learning. *Cognition and Instruction, 33*(3), 226–256. https://doi.org/10.1080/07370008.2015.1067215

Caldarella, P., Larsen, R. A. A., Williams, L., Downs, K. R., Wills, H. P., & Wehby, J. H. (2020). Effects of teachers' praise-to-reprimand ratios on elementary students' on-task behaviour. *Educational Psychology*, 1–17. https://doi.org/10.1080/01443410.2020.1711872

Carroll, J. B. (1963). A model of school learning. *Teachers College Record, 64*(8), 723–733.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*(1), 5–51. https://doi.org/10.1177/1529100618772271

CESE (Centre for Education Statistics and Evaluation). (2017). *Cognitive load theory: Research that teachers really need to understand Centre for Education Statistics and Evaluation*. https://www.cese.nsw.gov.au/images/stories/PDF/cognitive-load-theory-VR_AA3.pdf

Clausen, M. (2002). *Unterrichtsqualität: eine Frage der Perspektive? : empirische Analysen zur Übereinstimmung, Konstrukt- und Kritieriumsvalidität*. Waxmann.

Coe, R. (2014, January 9). *Classroom observation: It's harder than you think*. CEM Blog. http://www.cem.org/blog/414/

Coe, R., Aloisi, C., Higgins, S., & Major, L. E. (2014). *What makes great teaching? Review of the underpinning research*. Sutton Trust.

Coe, R. (1998). Can feedback improve teaching? A review of the social science literature with a view to identifying the conditions under which giving feedback to teachers will result in improved performance. *Research Papers in Education, 13*(1), 43–66. https://doi.org/10.1080/0267152980130104

Creemers, B. P. M., & Kyriakides, L. (2006). Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. In *School Effectiveness and School Improvement* (Vol. 17, Issue 3, pp. 347–366). Routledge. https://doi.org/10.1080/09243450600697242

Creemers, B. P. M., & Kyriakides, L. (2011). *Improving quality in education: Dynamic approaches to school improvement*. Routledge. https://doi.org/10.4324/9780203817537

Creemers, B., & Kyriakides, L. (2015). Process-product research: A cornerstone in educational effectiveness research. *The Journal of Classroom Interaction, 50*(2), 107–119. http://www.jstor.org/stable/44735492

Creemers, B., Kyriakides, L., & Antoniou, P. (2013). *Teacher professional development for improving quality of teaching*. Springer. https://doi.org/10.1007/978-94-007-5207-8

Csikszentmihalyi, M., & Schneider, B. (2000). Becoming adult: How teenagers prepare for the world of work. In *Becoming adult: How teenagers prepare for the world of work*. (pp. xx, 289–xx, 289). Basic Books.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Association for Supervision and Curriculum Development.

Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives, 8*(1), 1. https://doi.org/10.14507/epaa.v8n1.2000

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology, 49*(3), 182–185. https://doi.org/10.1037/a0012801

Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 53, Issue C, pp. 63–147). Academic Press. https://doi.org/10.1016/S0079-7421(10)53003-2

Department for Education. (2019). *Early Career Framework*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/773705/Early-Career_Framework.pdf

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process* (2nd ed.). D.C. Heath and Co.

Donker, A. S., de Boer, H., Kostons, D., Dignath van Ewijk, C. C., & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review, 11, 1–26*. https://doi.org/10.1016/j.edurev.2013.11.002

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques. *Psychological Science in the Public Interest, 14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.

Education Endowment Foundation (EEF). (2019). *Improving behaviour in schools*.

Education Endowment Foundation (EEF). (2018). *Metacognition and self-regulated learning: Guidance report*.

Education Endowment Foundation (EEF). (2020). *Improving mathematics in the Early Years and Key Stage 1*.

Ericsson, K. A. (2009). *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (K. A. Ericsson (ed.)). Cambridge University Press. https://doi.org/10.1017/CBO9780511609817

Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research, 11*(2), 147–252. https://doi.org/10.1016/0883-0355(87)90035-8

Good, T. L., & Brophy, J. E. (1984). *Looking in classrooms*. Harper & Row.

Gruehn, S. (2000). *Unterricht und schulisches Lernen : Schüler als Quellen der Unterrichtsbeschreibung*. Waxmann.

Guay, F., Ratelle, C. F., & Chanal, J. (2008). Optimal learning in optimal contexts: The role of self-determination in education. *Canadian Psychology/Psychologie Canadienne, 49*(3), 233–240. https://doi.org/10.1037/a0012758

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.). (2009). *Handbook of metacognition in education*. Routledge/Taylor & Francis Group.

Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children's development. *Child Development, 85*(3), 1257–1274. https://doi.org/10.1111/cdev.12184

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

Hattie, J., & Timperley, H. (2007). The Power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hill, H. C., & Charalambous, C. Y. (2012). Teacher knowledge, curriculum materials, and quality of instruction: Lessons learned and open issues. *Journal of Curriculum Studies, 44*(4), 559–576. https://doi.org/10.1080/00220272.2012.716978

Hill, H. C., & Chin, M. (2018). Connections between teachers' knowledge of students, instruction, and achievement outcomes. *American Educational Research Journal, 55*(5), 1076–1112. https://doi.org/10.3102/0002831218769614

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406. https://doi.org/10.3102/00028312042002371

Hogarth, R. M. (2001). *Educating intuition*. Chicago.

Hogarth, R. M., Lejarraga, T., & Soyer, E. (2015). The two settings of kind and wicked learning environments. *Current Directions in Psychological Science, 24*(5), 379–385. https://doi.org/10.1177/0963721415591878

Jones, S. M., & Doolittle, E. J. (2017). Social and emotional learning: Introducing the issue. *Future of Children, 27*(1), 3–12. https://doi.org/10.1353/foc.2017.0000

Kaiser, G., & König, J. (2019). Competence measurement in (mathematics) teacher education and beyond: Implications for policy. *Higher Education Policy, 32*(4), 597–615. https://doi.org/10.1057/s41307-019-00139-z

Kennedy, M. (2016). Parsing the Practice of Teaching. *Journal of Teacher Education, 67*(1), 6–17. https://doi.org/10.1177/0022487115614617

Kern, L., & Clemens, N. H. (2007). Antecedent strategies to promote appropriate classroom behavior. *Psychology in the Schools, 44*(1), 65–75. https://doi.org/10.1002/pits.20206

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254.

Ko, J., Sammons, P., & Bakkum, L. (2013). *Effective teaching: A review of research and evidence.* CfBT Education Trust. https://eric.ed.gov/?id=ED546794

Kounin, J. S. (1977). *Discipline and group management in classrooms.* R.E. Krieger Pub. Co.

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research, 88*(4), 547–588. https://doi.org/10.3102/0034654318759268

Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education, 36,* 143–152. https://doi.org/10.1016/j.tate.2013.07.010

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491. https://doi.org/10.3102/00028312032003465

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

London, M. (2003). *Job feedback: Giving, seeking, and using feedback for performance improvement.* Lawrence Erlbaum Associates.

Lynch, K., Hill, H. C., Gonzalez, K. E., & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. *Educational Evaluation and Policy Analysis, 41*(3), 260–293. https://doi.org/10.3102/0162373719849044

Metzler, J., & Woessmann, L. (2012). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation. *Journal of Development Economics, 99*(2), 486–496. https://doi.org/10.1016/j.jdeveco.2012.06.002

Meyer, J. H. F., & Land, R. (2005). Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. *Higher Education, 49*(3), 373–388. https://doi.org/10.1007/s10734-004-6779-5

Moore, D., Benham-Clarke, S., Kenchington, R., Boyle, C., Ford, T., Hayes, R., & Rogers, M. (2019). *Improving behaviour in schools: Evidence review*. https://educationendowmentfoundation.org.uk/public/files/Improving_B

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology, 75*(1), 33–52. https://doi.org/10.1037/0022-3514.75.1.33

Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art - teacher effectiveness and professional learning. *School Effectiveness and School Improvement, 25*(2), 231–256. https://doi.org/10.1080/09243453.2014.885451

Muijs, D., Reynolds, D., Sammons, P., Kyriakides, L., Creemers, B. P. M., & Teddlie, C. (2018). Assessing individual lessons using a generic teacher observation instrument: How useful is the International System for Teacher Observation and Feedback (ISTOF)? *ZDM - Mathematics Education, 50*(3), 395–406. https://doi.org/10.1007/s11858-018-0921-9

Nuthall, G. (2007). *The hidden lives of learners*. NZCER Press.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development*. https://doi.org/10.1159/000271225

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 365–386). Springer. https://doi.org/10.1007/978-1-4614-2018-7_17

Praetorius, A. K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of Three Basic Dimensions. *ZDM - Mathematics Education, 50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4

Rosenshine, B. (1976). Classroom instruction. In N. L. Gage (Ed.), *The psychology of teaching methods* (pp. 335–371). University of Chicago Press.

Rosenshine, B. (1987). Direct instruction. In M. J. Dunkin (Ed.), *International ecyclopedia of teaching and teacher evaluation* (pp. 257–262). Pergamon Press.

Rosenshine, B. (2010). Principles of instruction. *Educational Practices Series, 21*, 109–125. https://doi.org/10.1007/978-94-007-2669-7_7

Rosenshine, B., & Stevens, R. (1986). Teaching functions. *Handbook of Research on Teaching, January 1986*, 376–391.

Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Educational Research Journal, 50*(5), 1020–1049. https://doi.org/10.3102/0002831213477680

Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. Pergamon Press.

Scheerens, J., Luyten, H., Steen, R., & Luyten-de Thouars, Y. (2007). *Review and meta-analyses of school and teaching effectiveness*.

Schon, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454–499. https://doi.org/10.3102/0034654307310317

Shimamura, A. (2018). *MARGE. A whole-brain learning approach for students and teachers*. https://shimamurapubs.files.wordpress.com/2018/09/marge_shimamura.pdf

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. https://doi.org/10.3102/0013189X015002004

Smith, T. W., Baker, W. K., Hattie, J., & Bond, L. (2008). A validity study of the certification system of the National Board for Professional Teaching Standards. In R. E. Stake, S. Kushner, L. Ingvarson, & J. Hattie (Eds.), *Assessing teachers for professional certification: The first decade of the National Board for Professional Teaching Standards* (Vol. 11, pp. 345–378). Emerald Group Publishing Limited. https://doi.org/10.1016/S1474-7863(07)11012-7

Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science, 10*(2), 176–199. https://doi.org/10.1177/1745691615569000

Stenhouse, L. (1975). *An introduction to curriculum research and development*. Pearson Education.

Sternberg, R. J. (1988). Mental self-government: A theory of intellectual styles and their development. *Human Development, 31*(4), 197–224. https://doi.org/10.1159/000275810

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

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. In *Educational Psychology Review* (Vol. 31, Issue 2, pp. 261–292). Springer New York LLC. https://doi.org/10.1007/s10648-019-09465-5

Sweller, J., Van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development*. 344. http://educationcounts.edcentre.govt.nz/goto/BES

van de Grift, W. J. C. M., Chun, S., Maulana, R., Lee, O., & Helms-Lorenz, M. (2017). Measuring teaching quality and student engagement in South Korea and The Netherlands. *School Effectiveness and School Improvement, 28*(3), 337–349. https://doi.org/10.1080/09243453.2016.1263215

van Merriënboer, J. J. G., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology, 20*(3), 343–352. https://doi.org/10.1002/acp.1250

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology, 87*(2), 246–260. https://doi.org/10.1037/0022-3514.87.2.246

Vygotsky, L. S. (1962). *Thought and language* (E. Hanfmann & G. Vakar (Trans.)). MIT Press.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89–122. https://doi.org/10.3102/00346543073001089

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*(4), 548–573. https://doi.org/10.1037/0033-295X.92.4.548

Weinstein, Y., Sumeracki, M., & Caviglioli, O. (2018). *Understanding how we learn: A visual guide.* Routledge.

Wiliam, D. (2018). *Creating the schools our children need: Why what we're doing now won't help much (and what we can do instead).* Learning Scineces International.

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18–40). Routledge. https://doi.org/https://doi.org/10.4324/9780203874851

Willingham, D. T. (2019, July 14). *The high price of multitasking*. The New York Times. https://www.nytimes.com/2019/07/14/opinion/multitasking-brain.html

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*(2), 267–301. https://doi.org/10.3102/0034654311405999