Approaches to scaling innovations across schools: an analysis of key theories and models

Corresponding author:
Professor Toby Greany, University of Nottingham, Nottingham, UK.
Email: toby.greany@nottingham.ac.uk

Abstract

The question of how to scale up successful innovations remains a key challenge for policy and practice in education (Hall, 2013). This article outlines three established 'schools of thought' around how to scale up successful innovations. The first approach – eco-systems - is informed by complexity theory and tends to focus on generating and scaling innovations through diverse lateral networks which allow multiple stakeholders to collaborate (OECD, 2015). The second approach – 'what works?' – is interested in identifying effective practices through rigorous, scientific research and in replicating these proven practices across multiple schools with high degrees of fidelity. The third approach – Improvement Science - emphasises the importance of practitioner judgement and learning for interpreting and applying evidence across different contexts, but also recognises the importance of school-level cultures and wider factors, and so seeks to generate systematic forms of innovation, evaluation and improvement at scale (Bryk, 2015; Resnick, 2010). This article concludes by discussing the respective contributions of these models and the implications for policy, practice and research.

Introduction

The question of how to scale up and sustain successful innovations remains a key challenge for policy and practice in education. For example, Gene Hall (2013:265) explains, somewhat wearily, that:

> Over the last several decades the change ritual has become almost predictable. The process begins with identifying a particular problem or symptom that must be addressed… Next, a specific program, process or product… is selected. Then teachers, schools and districts go through the ceremony of launching the "new way." Materials are delivered to schools and teachers attend introductory sessions before the new school year begins. The implicit assumption of policymakers and system leaders… is that the "new way" is now in place. Within one to three years there may be an evaluation to see if test scores have indeed increased. All too often, the finding is one of "no significant difference" between the new way and the old way.

This picture of failed reforms remains dispiritingly common, despite significant efforts over the past few decades to understand the principles and processes that underpin scale up. These efforts have accelerated since the 1990s, when there was a recognition that the dominant diffusion approach in use at that time (i.e. through training and materials) was ineffective. Writing in 1996, Richard Elmore highlighted how rarely innovations impacted on the instructional core of teachers' classroom practice, arguing that this reflected "the absence of a practical theory that takes account of the institutional complexities that operate on changes in practice" (cited in Glennan et al, 2004:12).

Writing soon after the start of the new millennium, Glennan et al (2004:27) argued that a level of consensus was emerging, at least in the US, in thinking about scale up. However, this article outlines three quite different 'schools of thought' (or, broad approaches) on scale up which have emerged since that time, suggesting that consensus had not, in fact, been reached. The conclusion draws out and discusses the differences between them and identifies implications for policy, research and practice.

Definitions and key concepts
Before outlining the approaches to scale up it is helpful to define key terms.

The first term is 'innovation', which is most simply defined as 'doing things differently in order to do them better'. Innovations are certainly novel, at least in the context they are introduced, and can reflect new products, services, processes and/or organisational designs. In an education context, therefore, innovations could potentially encompass novel approaches within the existing school offer (for example, a new approach to pedagogy or the curriculum or to student engagement), or a new approach that extends the traditional school offer (for example, to help address emerging issues that young people face, such as online safety, childhood obesity or mental health issues).

Innovations should lead to measurable improvements in outcomes (i.e. 'doing things better'), but this is where it can be difficult to distinguish between 'innovation' and 'improvement', if both must lead to measurable improvements in outcomes. 'Improvement' can be defined as a planned movement from one state to another, where the impact on outcomes is measured solely through existing educational accountability metrics (i.e. usually standardised test scores or school inspection grades). By contrast, the impact of an 'innovation' might well include, but should also go beyond existing accountability metrics and outcomes, not least because these may not yet exist in many of the areas that require genuine innovation. This approach also helps to distinguish 'innovation' from 'change', where 'change' implies movement from one state to another, but without necessarily any impact on outcomes.

In practice, the texts cited in this article use the terms 'innovation', 'improvement' and 'change' almost interchangeably, so all three are used here. Nevertheless, it is helpful to keep these three definitions in mind when comparing the 'schools of thought' because this allows us to see that they can have slightly different objectives, as discussed in the conclusion.

Findings from research into innovation, improvement and change can help enrich our understanding of scale up. For example, Hall and Hord have analysed processes of change in education over several decades (Hall, 2013). Through this work they developed the Concerns Based Adoption Model, with its three dimensions: i) Stages of Concern assesses the emotional and affective aspect of change, showing how individuals move through seven distinct stages as they engage with a new approach; ii) Levels of Use evaluates what people are doing relative to an innovation, moving through seven levels of proficiency as they move from novice to expert in the new way of working; while iii) Innovation Configurations represents the possible operational forms of a change, capturing the ways in which the application of a procedure varies across different contexts.

These three dimensions have clear implications for how we understand scale up. For example, they highlight that: change is a process not an event; change is always a personal experience, requiring growth in confidence and competence and impacting on emotions, so will differ for each individual; an organization does not change until the individuals within it actually implement the new way; and innovations will usually be adapted – or 'mutate' - to meet the context in which they are applied.

The dimensions also help us to evaluate the 'schools of thought'. For example, with reference to the 'what works?' approach, Hall (2013: 272) highlights that: "In the traditional positivist research paradigm, the assumption is that there are two groups, treatment and control. It is a dichotomous view: users and non-users. The Levels of Use diagnostic… changes the perspective use/non-use to using".

Turning to the definition of 'scale up', this is also used interchangeably with concepts such as 'implementation' and 'replication' in the literature. These latter terms both suggest some level of imposition, the rolling out of a pre-defined approach. 'Scale up' can also have this connotation, but Coburn helps to position it in a more nuanced way by highlighting that successful scale up requires that the receiving site must take ownership of a new approach if it is to be sustained. This signals the importance of agency, autonomy and stakeholder engagement in a process of 'co-creation' as part of the scale-up process. Coburn's definition is as follows:

> Scaling up not only requires spread to additional sites, but also consequential change in classrooms, endurance over time, and a shift such that knowledge and authority for the reform is transferred from the external organization to teachers, schools, and districts. Thus, I propose a conceptualization of scale comprising four interrelated dimensions: depth, sustainability, spread, and shift in reform ownership (Coburn, 2003:4, cited in Glennan et al, 2004:29).

Networks, complexity and innovation eco-systems
The first 'school of thought' – eco-systems - is reflected in the OECD's work on innovative learning systems (OECD, 2015). This report reflects work across multiple countries and is particularly focussed on how education systems might need to change in order to develop more flexible and appropriate learning that equips young people for life and work in the 21st Century.

The thinking is informed by complexity theory. It rejects "excessively bureaucratic models", "mechanical policy metaphors" and "the assumption of central policy omnipotence within well-defined and controllable 'systems'", arguing instead that "more organic metaphors and models might seem messy and unpredictable, but eco-systems and complexity have become the nature of the contemporary world" (ibid, 17).

This focus on eco-systems seeks to recognise that multiple organisations and constituencies can influence learning, and that learning does not happen only in schools. Instead it sees a key role for the 'meso level', which sits above schools/learning providers and comprises various networks, communities, chains and initiatives. These networks can grow and diffuse innovations by promoting "horizontal connectedness across learning activities and subjects, in and out-of-school" (ibid, 18).

The assumption in this model is thus that innovations will emerge and spread organically as practitioners from diverse contexts interact with each other in networks and in a spirit of curiosity to address learner needs. Mechanistic models of 'scale up' are explicitly rejected (ibid, 81). Instead, the role of policy is to "set conditions and create climates" (ibid, 81), including by regulating, incentivising and accelerating change, for example by reducing standardisation and developing appropriate accountability metrics. Meanwhile, the role of system leaders in the meso layer is to: find examples of "positive deviance"; to build social capital by linking innovators together in peer-to-peer networks (such as Professional Learning Communities) so that they can learn from and with each other; and to inject knowledge, evidence and energy so that barriers to progress are unblocked.

'What works?': hard science and fidelity to evidence-based practices
The second 'school of thought' - 'what works?' – seeks to embed evidence-based practice. The rationale for the approach is that education should aim to be like medicine, where practice is underpinned by robust evidence drawn from scientific evaluations, such as Randomised Controlled Trials and other Quasi-Experimental Designs that use control and intervention groups to avoid selection bias. Thus, it is argued, teachers and schools should be required and supported to apply only proven practices that have been rigorously evaluated and shown to lead to measurable improvements in outcomes.

The leading proponents of this approach are Bob Slavin and Nancy Madden, who developed the Success for All[1] (SFA) school literacy programme in the US and the Institute for Effective Education[2] in the UK.

Slavin and Madden see scale up as primarily about ensuring fidelity to proven approaches. For example, they explain that "unlike many alternative schoolwide change models, SFA is not reinvented for each school staff… (because) we want to be sure that schools are implementing a form of the program that is true to the model that has been evaluated and found to be effective" (Glennan et al:140). They argue that a "long co-development process (would risk) losing the initial enthusiasm and readiness for change necessary for a staff to fully embrace a new schoolwide program" (ibid:140). SFA has a number of features that ensure this consistency, including highly trained facilitators who provide intensive training, together with tightly scripted curriculum materials for teachers. In order to secure buy-in from teachers, SFA requires that at least 80% of staff must vote in a secret ballot to adopt the approach.

Improvement science: learning to improve
Improvement Science has emerged relatively recently as a coherent approach to scale up, although it builds on long-standing traditions of practitioner engagement in evidence-informed improvement (for example through action research and enquiry networks) and on realist evaluation methods that ask 'what works, for whom, under what circumstances and why?' It also draws directly on successful practices in industry and the health sector.

---

[1] See https://www.successforall.org/ accessed 15.3.19
[2] See https://the-iee.org.uk/ accessed 15.3.19

Improvement Science recognises that organisations are complex and so assumes that teachers and schools must be individually and collectively engaged in a continual process of learning how to improve, thereby developing 'practice-based evidence'. This learning is structured through cycles of improvement that are designed to develop, test, and refine interventions aimed at addressing specific problems.

The Carnegie Foundation for the Advancement of Teaching,[3] in the US, has been integral in promoting Improvement Science in education, which it describes in six steps:
1. Make the work problem-specific and user-centered, starting with the question: "What specifically is the problem we are trying to solve?"
2. Variation in performance is the core problem to address, so the aim should be to help everyone learn together how to improve at scale.
3. See the system that produces the current outcomes. Go and see how local conditions shape work processes. Make your hypotheses for change public and clear.
4. We cannot improve at scale what we cannot measure. Embed measures of key outcomes and processes to track. Anticipate unintended consequences and measure these too.
5. Anchor practice improvement in disciplined inquiry. Engage in rapid cycles of Plan, Do, Study, Act (PDSA) to learn fast, fail fast, and improve quickly.
6. Accelerate and broaden improvements through networked communities.

Tony Bryk is arguably the key proponent of Improvement Science in education, in his role as President of the Carnegie Foundation, although Lauren Resnick's (2010) work on organisational routines and the Hall and Hord work on change referenced above have informed the approach. Bryk (2015:473) positions Improvement Science as a direct alternative to 'what works?', arguing that: "randomized field trials are not principally designed to tell us how to make interventions work reliably for different subgroups of students and teachers working under varying contextual conditions. Improvement Science, in contrast, places primacy on variability in outcomes as the central problem to address. It focuses our attention on how task and organization factors combine to create this variability."

Improvement Science emphasises the importance of peer networks, and is clearly focussed on strengthening the professionalism of teachers, so has some similarities to the eco-system model. However, Improvement Science is more deliberate and disciplined, with practitioners engaging with researchers and others to develop, test, and enhance the clinical work of schooling, with common measures, inquiry methods, and communication mechanisms to anchor collective problem solving.

Conclusion
These short overviews highlight significant differences between the three 'schools of thought', at both philosophical and practical levels. The emerging consensus around how to achieve scale up identified by Glennan et al in 2004 has not, in fact, emerged. Meanwhile, the pace of social, technological, environmental and economic change has continued to increase since Glennan et al were writing. These changes often impact directly on the lives of children and on the work of schools, increasing the need for innovation at scale.

---

[3] See https://www.carnegiefoundation.org/ accessed 15.3.19

How then should we understand the differences between the three 'schools of thought' and what are the implications for policy, practice and research? In making such an assessment we must remember the differences between 'improvement' and 'innovation' outlined above: if we only use existing measures of school effectiveness (such as improvements in standardised pupil test scores) as proxies for effective scale up, we risk elevating 'improvement' over 'innovation', because the latter seeks to be more responsive to real needs and to go beyond existing accountability metrics.

One way to compare the three 'schools' is to consider how far they reflect the conclusions reached by Glennan et al, who argued that:

a the scale-up process is necessarily iterative and complex and requires the support of multiple actors;

b the actors must jointly address a set of known, interconnected tasks if scale-up is to succeed; and

c four factors influence the success of scale up efforts - the characteristics of the intervention, school conditions, assistance provided for implementation, and alignment of the policy and infrastructure supports.

A simple assessment might be that: the 'eco-systems' approach addresses point a) most securely; 'Improvement Science' addresses b); and 'what works' addresses c). The eco-system approach is focussed on 'innovation': it prioritises professional autonomy and assumes that committed professionals will automatically develop and adopt innovative approaches that respond to learner needs if the conditions are right. The 'what works?' approach is interested in 'improvement', through the application of tried and tested approaches. It positions teachers as clinical technicians, trained to apply proven practices through a policy-controlled process. Improvement Science sits somewhere between these two models: it is focussed more on improvement than innovation, but it attempts to do this through processes which activate teacher agency in disciplined and evidence-informed ways.

So how have the three approaches been applied and what is the evidence of their success in achieving scale up? What are the main critiques in each case?

The 'what works?' approach has been the most influential, particularly in the US and England, where huge amounts of money have been spent on evaluating and scaling up interventions. For example, in the US, the Investing in Innovation (i3) grant program[4] was established in 2009 to fund the development, evaluation, and scale-up of proven programs. By its final year, in 2016, i3 had funded 171 projects and spent more than $1.4 billion. Some, though by no means all, of the interventions that have been scaled up through this process have been evaluated to show impact, including Success for All. However, critics of 'what works' approaches argue that they are slow, expensive and overly instrumental, and that they fail to reflect the importance of context (as expressed in the quote from Bryk above). Equally, the US and England, where 'what works?' has been applied most

---

[4] See https://www.ed.gov/open/plan/investing-innovation-i3 accessed 15.3.19. Similarly, in the UK, the Education Endowment Foundation (EEF) received £125 million from government in 2010 and has funded more than 107, mostly randomized evaluations of programs.

extensively, are by no means high performers in international assessments such as PISA, suggesting that it is far from a panacea.

The eco-systems approach is harder to assess, either in terms of how far it has been applied and whether it has been successful in achieving scale up, not least because it is deliberately organic and therefore hard to quantify or evaluate. The OECD report does include examples from different countries to illustrate the approach, although it is clear that these are not proven 'best practices' (ibid, 5). Certainly, many school systems around the world are working to foster lateral school networks and to encourage teachers to collaborate in PLCs, but evidence of impact from these initiatives remains thin. Bryk critiques PLCs on the basis that they depend "heavily on individual educators' tacit knowledge… (with) no formal mechanism for accumulating, further detailing, and testing this individual clinical knowledge so that it might be transformed over time into a collectively held professional knowledge" (ibid, 469).

Finally, the Improvement Science approach is relatively recent, and so has not been widely applied and evaluated, although it does have some early evidence of success. For example, Resnick (2010) describes evaluations of approaches in US schools that reflect many of the principles of Improvement Science and which have had a positive impact. Improvement Science does appear to address the lessons from research in this area, for example in its consideration of organisational climate issues and its focus on building collective efficacy. However, Improvement Science remains relatively untested and may prove to have its own pitfalls. For example, it assumes that improvement outcomes will always be collectively owned and understood across an organisation, but such clarity may only occur in the context of high stakes accountability systems which, by their nature, limit the scope for innovations that go beyond measurable improvement.

In conclusion, despite the lack of consensus around how to scale up successful innovations, there are some clear implications for policy, practice and research emerging. The fact that there are (at least) three 'schools of thought', each with its own theoretical assumptions and associated strengths and weaknesses, could be seen as helpful, because it gives policy makers and practitioners choice. This article has sought to highlight that these choices are not simply about the relative effectiveness of the different approaches. Rather, it is about considering context and aims: if the focus is on 'improvement', then 'what works?' or Improvement Science may be most appropriate, but if the focus is on 'innovation', then the eco-systems model may be more suitable. Similarly, from a research perspective, there is an urgent need to strengthen our understanding of when and how schools and educational systems do actually achieve successful innovation at scale.

References:

A. Bryk. Accelerating How We Learn to Improve. *Educational Researcher*, 44/9, pp. 467–477 (2015) DOI: 10.3102/0013189X15621543

T. K. Glennan Jr. S. J. Bodilly, J. R. Galegher, K. A. Kerr (Eds). *Expanding the Reach of Education Reforms: Perspectives from Leaders in the Scale-Up of Educational Interventions* (Santa Monika, CA: RAND Corporation, 2004).

G. E. Hall, Evaluating change processes: Assessing extent of implementation (constructs, methods and implications). *Journal of Educational Administration*, 51/3, pp.264-289 (2013)

OECD (2015), *Schooling Redesigned: Towards Innovative Learning Systems*, Educational Research and Innovation, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264245914-en

L. B. Resnick, Nested Learning Systems for the Thinking Curriculum. *Educational Researcher*, 39/3, pp. 183–197 (2010). DOI: 10.3102/0013189X10364671