

Building to Impact

Do not copy, post, or distribute

Do not copy, post, or distribute

Building to Impact

The 5D Implementation Playbook for Educators

Arran Hamilton

Douglas B. Reeves

Janet M. Clinton

John Hattie

Foreword by Dylan Wiliam

CORWIN

Copyright ©2022 by SAGE Publications, Inc.

This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

FOR INFORMATION:

Corwin
A SAGE Company
2455 Teller Road
Thousand Oaks, California 91320
(800) 233-9936
www.corwin.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
18 Cross Street #10-10/11/12
China Square Central
Singapore 048423

President: Mike Soules
Associate Vice President and
Editorial Director: Monica Eckman
Publisher: Jessica Allan
Senior Content
Development Editor: Lucas Schleicher
Content Development Editor: Mia Rodriguez
Editorial Assistant: Natalie Delpino
Project Editor: Amy Schroller
Copy Editor: Christina West
Typesetter: C&M Digitals (P) Ltd.
Cover Designer: Scott Van Atta
Marketing Managers: Deena Meyer and
Olivia Bartlett

Copyright © 2022 by Corwin

All rights reserved. Except as permitted by U.S. copyright law, no part of this work may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without permission in writing from the publisher.

When forms and sample documents appearing in this work are intended for reproduction, they will be marked as such. Reproduction of their use is authorized for educational use by educators, local school sites, and/or noncommercial or nonprofit entities that have purchased the book.

All third-party trademarks referenced or depicted herein are included solely for the purpose of illustration and are the property of their respective owners. Reference to these trademarks in no way indicates any relationship with, or endorsement by, the trademark owner.

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Names: Hamilton, Arran, author. | Reeves, Douglas B., 1953- author. | Clinton, Janet M. (Janet May), author. | Hattie, John, author.

Title: Building to impact : the 5D implementation playbook for educators / Arran Hamilton, Douglas B. Reeves, Janet May Clinton, John Hattie.

Description: First edition. | Thousand Oaks, California : Corwin, [2022] | Includes bibliographical references and index.

Identifiers: LCCN 2022004775 | ISBN 9781071880753 (paperback) | ISBN 9781071880760 (epub) | ISBN 9781071880777 (epub) | ISBN 9781071880784 (pdf)

Subjects: LCSH: School improvement programs. | Educational innovations.

Classification: LCC LB2822.8 .H34 2022 | DDC 371.2/07—dc23/eng/20220310

LC record available at <https://lccn.loc.gov/2022004775>

This book is printed on acid-free paper.

22 23 24 25 26 10 9 8 7 6 5 4 3 2 1

DISCLAIMER: This book may direct you to access third-party content via Web links, QR codes, or other scannable technologies, which are provided for your reference by the author(s). Corwin makes no guarantee that such third-party content will be available for your use and encourages you to review the terms and conditions of such third-party content. Corwin takes no responsibility and assumes no liability for your use of any third-party content, nor does Corwin approve, sponsor, endorse, verify, or certify such third-party content.

Contents

List of Figures	vii
Foreword	xi
Dylan Wiliam	
Preface	xvii
About the Authors	xxiii
INTRODUCTION	1
D1 DISCOVER	19
D2 DESIGN	49
D3 DELIVER	101
D4 DOUBLE-BACK	125
D5 DOUBLE-UP	149
CONCLUSION	165
Appendices	
Appendix 1: Summary of 50 Reviewed Implementation Models	171
Appendix 2: Key Areas of Divergence Across 50 Implementation Models	179
Appendix 3: Key Areas of Similarity Across 50 Implementation Methodologies	183
Appendix 4: Findings From Selected Meta-Analyses and Systematic Reviews of Implementation	187
Glossary	189
References	193
Index	203

Do not copy, post, or distribute

List of Figures

Figure 0.1:	Percentage of Students Achieving the Minimum Threshold on International Assessments, 1970–2015	2
Figure 0.2:	Percentage of Students Achieving the Advanced Threshold on International Assessments, 1970–2015	3
Figure 0.3:	<i>Building to Impact 5D</i>	11
Figure 0.4:	Unpacking the <i>Building to Impact 5D</i> Stages	12
Figure 0.5:	A Walk-Through of <i>Building to Impact 5D</i>	14
Figure 0.6:	Iterative Implementation	15
Figure 0.7:	The Double-Back Mind Frame	15
Figure 0.8:	Systematic Reviews on Implementation	16
Figure 1.1:	Example Responsibilities and Accountabilities Matrix	24
Figure 1.2:	Sources of Data	31
Figure 1.3:	The Education Challenge Funnel	33
Figure 1.4:	Education Challenge Hunting	35
Figure 1.5:	Matched Pairs Comparison	35
Figure 1.6:	So What and So Prove It	36
Figure 1.7:	Education Challenge Breakdown Structure	39
Figure 1.8:	The <i>Five Whys</i> Approach	41
Figure 1.9:	Path Analysis	42
Figure 1.10:	Fishbone and Issues Tree Mapping	44
Figure 1.11:	Validating the Causal Drivers	45
Figure 1.12:	Establishing Provisional Success Criteria	47
Figure 2.1:	Recapping the Path Analysis	52
Figure 2.2:	Identification of Options in the Design Space	54
Figure 2.3:	Identification of Positive Outliers	55
Figure 2.4:	Global Education Research Repositories	57

Figure 2.5:	Not All Evidence Is the Same	58
Figure 2.6:	Evidence to Programs vs. Programs to Evidence	58
Figure 2.7:	Opportunity Sketch Mapping	59
Figure 2.8:	Ranking Your Options in the Design Space	61
Figure 2.9:	Example Opportunity Sketch Analysis Framework	63
Figure 2.10:	A Rube Goldberg Machine	64
Figure 2.11:	Recapping Key Prior Steps	68
Figure 2.12:	Program Logic Model Template	69
Figure 2.13:	Partially Completed Program Logic Model	70
Figure 2.14:	Design Features and Setting Levels	74
Figure 2.15:	The Graphic Equalizer	76
Figure 2.16:	Theory 1 vs. Theory 2	78
Figure 2.17:	Program Logic Model Risk Mitigation Planning	82
Figure 2.18:	The STOP Audit	84
Figure 2.19:	The Cognitive Bias Codex	84
Figure 2.20:	Program Logic Model—STOPPING and STARTING	86
Figure 2.21:	Program Logic Model—Example of Stopping and Starting	87
Figure 2.22:	Purpose: Monitoring vs. Evaluating	90
Figure 2.23:	Black-Box, Gray-Box, and Clear-Box Evaluative Approaches	91
Figure 2.24:	Levels of Evaluation	92
Figure 2.25:	Level of Adoption	93
Figure 2.26:	Evaluative Indicator Selection— Worked Example	95
Figure 2.27:	Evaluation Plan Methods Grid	96
Figure 2.28:	Six Approaches to Target Setting	97
Figure 2.29:	Target Setting—Worked Example	98
Figure 2.30:	Monitoring and Evaluation in the Program Logic Model	99
Figure 3.1:	Delivery Approach Options Basket	104
Figure 3.2:	Common Project Management Assets and Tools	108
Figure 3.3:	Gantt Chart Illustration	109

Figure 3.4:	Kanban Board Illustration	110
Figure 3.5:	Highlight Report Illustration	112
Figure 4.1:	St. Louis to Los Angeles, 1804	126
Figure 4.2:	Easier, Hard, and Super-Hard Education Challenges	128
Figure 4.3:	Embedded Evaluation Processes (i.e., Double-Back)	129
Figure 4.4:	Zombie vs. Real Evaluation	130
Figure 4.5:	Recapping the Program Logic Model	131
Figure 4.6:	Time Tracking	133
Figure 4.7:	Recapping the Six Levels of Evaluation	135
Figure 4.8:	Level of Engagement	137
Figure 4.9:	Iterative Implementation	138
Figure 4.10:	Four-Quadrant Matrix	139
Figure 4.11:	Adjusting the Design Features and Setting Levels	141
Figure 4.12:	An Example of Program Logic Model Implementation, Iteration, and STOPPING	143
Figure 4.13:	Three Lenses for Evaluating Your Evaluation	144
Figure 4.14:	Generating Perverse Incentives	145
Figure 4.15:	Meta-Evaluative Questions	146
Figure 4.16:	Iterative Adaptation Across Time	147
Figure 5.1:	Different Models for Scaling	157
Figure 5.2:	The Scaling Checklist	162

Do not copy, post, or distribute

Foreword

*From *schools don't make a difference (wrong), to *schools (can) make a big difference, to *schools aren't (yet) making enough of a difference, to *school improvement requires systematic and highly exacting implementation*

For much—perhaps most—of the last century, educational policy in many countries was underpinned by two assumptions. The first was that “intellectual capacity was *wholly* or *very largely* determined by genetic endowment and was therefore fixed, unchanging, and, in addition, accurately measurable by group intelligence tests” (Chitty, 1997). The purpose of schools was therefore not about “providing an enriching and creative environment, but should be adjusted to the function of sorting out and selecting the ‘bright’ from the ‘dull,’ as determined by nature, and as basically reflected in the existing social hierarchy” (Lawler, 1978, p. 3).

While some such as Arthur Jensen continued to argue that general cognitive ability was inherited, others such as Basil Bernstein argued that the child’s environment mattered far more. Either way, many believed that students’ academic achievement was largely predetermined and schools could do little to change this. As Bernstein (1970) put it, “Education cannot compensate for society.”

The second assumption was that the quality of schooling had little impact on student achievement. In other words, as long as students attended school, which school they attended had little impact on how much they learned. As Chitty (1997) notes, Jensen and Bernstein undoubtedly contributed to the general mood of “educational fatalism” that pervaded policy debates in England and in the United States in the early 1970s, but perceptions about the influence of schools on educational achievement were also strongly supported by two large-scale surveys conducted in the United States (p. 49).

The Equality and Educational Opportunity Study (Coleman, 1966) surveyed over 600,000 first-, third-, sixth-, ninth-, and twelfth-grade students in 3,000 US elementary and secondary schools. While the original remit of the study was to describe the level of inequality across US schools by focusing on “inputs,” the study team broadened the scope considerably, collected information about the attitudes of teachers and administrators, and assessed students using standardized tests of ability and achievement.

The report's conclusion was stark:

Taking all these results together, one implication stands out above all: Schools bring little influence to bear on a child's achievement that is independent of his background and general social context; and that this very lack of an independent effect means that the inequalities imposed on children by their home, neighborhood, and peer environment are carried along to become the inequalities with which they confront adult life at the end of school. (Coleman, 1966, p. 325)

Six years later, Christopher Jencks and his colleagues reanalyzed the data from the Coleman report together with data from a range of other sources and came to a similar conclusion that “equalizing the quality of high schools would reduce cognitive inequality by one per cent or less” (Jencks et al., 1972, p. 109).

While many academics, policymakers, and educators seemed happy to accept such strong claims, others pointed out that the datasets analyzed by Coleman and Jencks were cross-sectional rather than longitudinal, and therefore cast little light on the progress made by students in school. Even with better data, if it was shown that some students did not make much progress in schools, this would tell us little about what might be, not least because many practices in schools—such as grouping students by ability—were based on the assumption that ability was largely fixed. Therefore, the observed results were likely to be as much a product of pre-existing assumptions about students than about what schools might be able to achieve.

The idea that schools *could* make a difference received support from early international comparisons of student achievement, such as the First International Mathematics Study (Husén, 1967a, 1967b). A special issue of *Forum* magazine in 1974 directly challenged Basil Bernstein's earlier claim with the title “Schools can make a difference” (Simon & Whitbread, 1974).

Strong empirical support for the idea that schools did in fact make a difference—and that some schools were significantly more effective than others—came toward the end of the 1970s. A study of eight schools in Michigan by Brookover and Lezotte (1977) found clear differences between six schools where student performance was increasing and two where performance was declining. In the more successful schools, teachers believed that all students could master basic objectives, were less satisfied with their achievements, and held higher and increasing expectations of their students (Edmonds, 1979).

Perhaps even more influential was the publication of *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children* by

Rutter and colleagues in 1979. Beginning in 1970 with a convenience sample of 10-year-olds in London who had taken nonverbal reasoning and reading tests in the final year of primary schools, the research team tracked approximately two-thirds of these children into secondary schools and followed their progress until they reached the end of compulsory schooling 5 years later. While the study report produced a range of useful findings, perhaps the most important was the conclusion about the impact of schools on academic achievement:

[T]he differences between schools in outcome were systematically related to their characteristics as social institutions. Factors as varied as the degree of academic emphasis, teacher actions in lessons, the availability of incentives and rewards, good conditions for pupils, and the extent to which children were able to take responsibility were all significantly associated with outcome differences between schools. All of these factors were open to modification by the staff, rather than fixed by external constraints. (Rutter et al., 1979, p. 178)

In the three decades following the publication of *Fifteen Thousand Hours*, as more and better data on school effects have become available, it has become widely accepted that schools differ in their effectiveness and, more importantly, that the factors that cause these differences are amenable to change. That said, there remains considerable disagreement about how much schools differ in their effectiveness, the factors that affect school effectiveness, and how schools can be supported in becoming more effective.

For example, while differences in student outcomes vary greatly from school to school, much of these differences appears to be attributable to factors outside the school's control. For many years, in addition to publishing the average results of students on England's national school-leaving examination, the General Certificate of Secondary Education (GCSE), for each school, the UK government also published a measure of "contextual value added" (CVA) that took account of the prior achievement of the students attending that school together with demographic characteristics (gender, ethnicity, socioeconomic status). For the 4,158 schools that had students taking GCSE examinations in 2007, the correlation between the average GCSE grades and the CVA measure was 0.27, suggesting that only around 8% of the variance in average student outcomes is currently attributable to the school (William, 2010). Analysis of data from the 2006 cycle of the Programme for International Student Assessment (PISA) examining variance in achievement not explained by the PISA index of economic, social, and cultural status of schools yields similar estimates for many countries:

Australia	7%
Denmark	6%
Finland	4%
New Zealand	4%
Spain	6%
Sweden	6%
United Kingdom	7%
United States	8%

Now, it is important to note that these figures tell us nothing about how good the education systems in different jurisdictions are. If all the schools are uniformly excellent, the achievement of students will be high, but the variation between schools will be small—as long as you go to school, it won't matter much which school you go to. However, these figures do tell us that similar students fare differently in different schools within the same system. More importantly, while school effects might only be 4–8% of the variation in student achievement on average, they can be hugely important for individual students.

Perhaps even more importantly, recent work on teacher quality suggests that one of the reasons that the variation between schools is small is because good teachers are fairly randomly distributed within the system, not least because it is rather difficult to identify more effective teachers with any accuracy (Wiliam, 2016). Schools can become much more effective if they support their teachers in improving.

In addition to data on the relative effectiveness of schools, PISA and other programs such as Trends in Mathematics and Science Study (TIMSS) and Progress in International Reading Study (PIRLS) have provided a wealth of information about the contextual factors influencing school performance. With all of this information about what makes schools more or less effective and with the undoubted desire of everyone working in education to make things better, we should have seen dramatic improvement in educational outcomes around the world. But we haven't.

Obviously, many—and heroic—assumptions are needed to compare results across the different cycles of PISA. But given the efforts that are made to secure the comparability of scores over time, the broad-brush conclusion from the seven cycles conducted so far (starting in 2007) is that—despite widespread, and often well-funded efforts to improve student achievement—there has been little net improvement in student achievement across the rich countries of the world.

The main reason is not, as might be assumed, that we don't know what to do. As the authors of this book point out, "We have more evidence about *what works best* for student outcomes than at any time in human history" (p. xx). As writers like Michael Fullan, Tony Bryk, Marc Tucker, and many others have pointed out, the reason is rather that most attempts to improve education have failed to address the simple fact that education systems are just that; they are *systems*. Knowing what to do is of no use if you can't implement your findings. Worse, changing one part of a system is likely to have little impact if the benefits of that change are offset by consequent changes in other parts of the system.

And this is why *Building to Impact* represents a step-change in thinking about educational improvement. Policymakers, leaders, and teachers—indeed, all of us—are, unsurprisingly, drawn to simple solutions. But as H. L. Mencken (1917/1949) wrote over 100 years ago, "there is always an easy solution to every human problem—neat, plausible, and wrong" (p. 443). Rather than pretending that school improvement is straightforward and easy, the authors face head on the inevitable complexity of improving schools. Drawing on a vast range of resources—implementation methodologies and processes, systematic reviews of research, and decades of "on the ground" experience working with teachers, leaders, and administrators—the authors present a rigorous approach that I believe represents the state of the art in school improvement.

As the authors themselves acknowledge, the 5 stages and the 18 key processes of *Building to Impact* are complex and can be challenging to implement, but that complexity should be a sign of their authenticity. After all, if there were easy solutions, we would have found them by now.

The work is not easy. As Seymour Sarason (1995) has pointed out,

The decision to undertake change more often than not is accompanied by a kind of optimism and rosy view of the future that, temporarily at least, obscures the predictable turmoil ahead. But that turmoil cannot be avoided and how well it is coped with separates the boys from the men, the girls from the women. It is . . . rough stuff . . . it has no end point, it is a continuous process, there are breakthroughs, but also brick walls. And it is indisputably worthwhile. (p. vii)

Right now, I know of no better guide to this process than the book you have in your hands.

—Dylan William
UCL Institute of Education

Do not copy, post, or distribute

Preface

*You've got to be very careful if you don't know where you are going,
because you might not get there*

Yogi Berra (1998)

Education takes you places.

The relationship between access to quality schooling and longer life expectancy, greater happiness, greater earnings, and a whole host of other good things is long established and rarely disputed (Hamilton & Hattie, 2022; World Bank, 2018). The reverse leads to the reverse. A two-way street.

It's brilliant, then, that governments should invest so much in education. We deploy a global army of more than 84 million teachers to take children places. The effect of this investment has also been relatively strong. In 1920, only 32% of the world was literate; over the following 100 years, this has risen to 86% (Roser & Ortiz-Ospina, 2016). Surely this is one of the greatest achievements of human history, greater than the Pyramids and all the other wonders combined. Maybe the greatest of all.

So why the need for this book?

Because it's still not enough. Despite the global investments in education, the returns are remarkably lumpy—among countries, between schools in the same country, and even teachers in adjacent classrooms in the same school! In the predominantly English-speaking world (e.g., the United States, United Kingdom, Canada, Australia, and New Zealand), performance on international student assessments is only moderately better today than in 1970. Yes, things are improving. But the incline is too shallow, and with ever larger financial investment and increasingly overworked teachers.

Often the proposed answer to this question of educational lumpiness is more research and the design of shiny new programs to close the equity gap. However, we are going to speak sacrilege: maybe there is already enough research to be getting on with?

By our count, if you wanted to consume *all* of the existing publications on education improvement, you would need to work through 68 books and journal articles *per day* to do it in a single lifetime. This wouldn't give you enough spare time to put any of it into practice, and the research continues to grow by the hour. There are also already thousands of programs and protocols based on the existing evidence, many of which have extremely strong evidence of impact. They really do shine.

We think the era of evidence collection is (largely) over. It's the era of systematic implementation of the existing evidence that now needs to begin.

Ergo, we think the era of evidence collection is (largely) over. It's the era of **systematic implementation of the existing evidence** that now needs to begin. In other words, what we really need most are good implementation processes and procedures. We need processes that enable systems, schools, and teaching teams to systematically discover their most pressing needs; to systematically select, localize, or design locally appropriate high-impact approaches, based on that vast existing evidence on effective shiny things; to then deliver to the designs systematically; and to then double-back (i.e., evaluate) in order to double-up (i.e., sustain and scale thy impact).

The idea is that these processes help schools and systems to bring the (truckloads of) pre-existing evidence to life, in a manner that supports locally relevant improvement, with locally available resources and local capabilities. For the avoidance of all doubt, this also includes local evidence collection in order to check that there was indeed impact.

Most of the existing research on effective implementation comes from outside education. There is a long history of it in industry, going back more than 100 years to the work of Frederick Winslow Taylor, Henry Gantt, and James O. McKinsey. Much of it is scientific—not in terms of generating complex formulas, but in how the thinking is done (i.e., in a stepwise fashion, with checking and cross-checking against data at each stage). It's systematic to the core. Albeit the system is applied to different contexts, with different needs and that it inevitably (and desirably) leads to different initiatives and different (positive) outcomes. Same, same but different.

Despite the growing evidence that the approach to implementation (i.e., whether you have an approach, whether it is “good,” and whether you actually follow it) is a major predictor of whether your improvement agenda is successful. Education has been a relative latecomer to the party. The first major handbook on education implementation did not hit the shelves until 2012 (i.e., Kelly & Perkins, 2012). Many more works have since been produced and this is strongly welcomed by us. However, they seem to fall into two broad camps:

- **Camp 1:** academic texts that wax lyrical on implementation process theory but that have limited practical application to schools and systems. Another variant of this camp is written in Malcolm Gladwell-like delicious speak. This may be a good (and fun) summer read but is challenging to put into action without developing your own process map and tools.
- **Camp 2:** practical primers that showcase some tools and approaches but not in enough detail that you could pick up the book and run with it. You still need to hire in the writer(s) or their teams to tell you the hidden extras or conjoin their processes with another body of knowledge.

What we felt was missing was a practical (but rigorous) step-wise process that takes educators all the way from discovering a pressing need to evaluating the impact of their selected high-probability interventions, and one that is crammed to the rafters with tools needed to support implementation and not just a couple of samples. So, we thought, why not write *that* book? And we did. And now you have it!

Building to Impact 5D is explicitly designed to be a field manual or playbook, something that you can pick up and use end to end or adapt to help you implement with rigor and impact. It combines the following:

- **Processes and practical tools** that we four (and our respective teams) already employ in our design, delivery, and evaluation work with systems, districts, and schools; and
- **A wide sweep of the global implementation science literature**, including a review of 50 implementation models and their respective tools as well as analysis of the available systematic reviews and meta-analysis on implementation success factors.

We think you will find our 5D playbook most useful if you are working at a school system, school district (i.e., overseeing several schools), schoolwide, or within-school level, seeking to generate sustainable improvement at scale. If you are part of a teaching team or a professional learning community, there's also plenty within for you to draw from. But you may not have the time or resources to follow the process from A to Z, so we also offer guidance on which elements you can undertake in a lighter-touch way.

We are not at all precious about the specifics of how you undertake local inquiry and implementation activity using the *Building to Impact 5D* playbook. You obviously need to localize to your context, resourcing level, culture, and time constraints. What we are more precious about is that you have to stop and ask each of the questions detailed within the framework before deciding what to do next—and the answers to those questions need to be generated by far more than a hunch, instinct, intuition, or gut feeling. You need to find and use data, and you need to explicitly look for disconfirming data, not just selective facts that (conveniently?) fit your preintended course of action. We are very strict about this.

As you will soon discover, there are 18 separate questions and/or processes you need to work through to properly implement *Building to Impact 5D* but you can flex the duration of each to your context. This could mean that a district-level team might spend several weeks digging and exploring one or two of these questions and/or process areas, whereas a school-level team might push forward

much sooner (e.g., after a quick brainstorming session), provided everyone comes with their data! You might also move back and forth between different steps and questions, reconsidering earlier decisions as new information comes to light. This is to be expected and is completely normal. We do it all the time.

It takes great resolve and discipline to think and act systematically and not slip back into intuitive, hunch-based ways of operating.

Despite this flexibility, the systems, districts, and schools that we work with often say that “it’s really hard to continuously work like this.” Amen to that. It takes *great* resolve and discipline to think and act systematically and not slip back into intuitive, hunch-based ways of operating. Effective implementation isn’t easy. The deliberateness of thought is the important part, but it is also the cognitively fatiguing bit. It really does make your head hurt. That’s why we think there is perhaps a missing role in schools and school systems: the implementation specialist or implementation scientist. This is someone whose sole job (or main job) is to support their colleagues with the mental heavy lifting and who is deeply trained in these processes. You might think of them as a cross between an educational strategy consultant, a project manager, and a formative evaluator—all in one. Other forms of specialists already exist in our world: data specialists, governance specialists, leadership improvement specialists, assessment specialists, and so on. So, why not implementation specialists? Think about it.

And now to the credits. Many people have supported and enhanced our thinking as we codified the *Building to Impact* 5D framework. These include Arran’s colleagues (past and present) at Cognition Education: Shaun Hawthorne, Mary Sinclair, Brian Hinchco, Phil Coogan, Mel Sproston, Helen Butler, Lindsey Conner, Jenna Crawley, Christophe Mullins, Nigel Bowen, Durgesh Rajandiran, and Tina Lucas.

John thanks all the consultants and implementers within the Corwin Visible Learning team who have taught him about implementation and their enablers and barriers, and he thanks the team within the Australian Institute for Teaching and School Leadership (AITSL) who show him so often what it means to translate evidence into action.

Janet thanks her evaluation team at the Centre for Program Evaluation and the leadership team at the Melbourne Graduate School of Education where so much implementation occurs.

Doug thanks his colleagues at Creative Leadership Solutions, Arran for orchestrating the book, and John for his continued global impact on education.

We all thank Dylan Wiliam, who took the time to go through our manuscript to help us buttress and improve and who also wrote the foreword! And the brilliant team at Corwin who brought

this project to life, especially Jessica Allan, Lucas Schleicher, Amy Schroller, and Christina West.

Obviously, our thinking didn't emerge from the ether. It very much builds on and from the heavy lifting done by others, including Michael Fullan, Thomas Guskey, Abraham Wandersman, Sir Michael Barber, Russell Bishop, Viviane Robinson, Lant Pritchett, and (the aforementioned) Dylan Wiliam. We owe these folks a great intellectual debt. They may not agree with all in this book but their words and actions have been a major inspiration.

And now back to the start.

In addition to being one of the greatest catchers in baseball history, Yogi Berra was something of a sage, renowned for his "Yogiisms." When he said, "You've got to be very careful if you don't know where you are going, because you might not get there," he was 10,000% right. And that's the whole point of *Building to Impact 5D*: to help you systematically decide on the destination, to then explore the different ways you could undertake the journey, to put one or more of those journey plans into action, to check whether it is working, and then to decide what to do next.

When you know where you are going, when you have a plan, and when you systematically check and revise, you will get there—albeit with unexpected twists and turns along the way.

If you want business as usual, move along. But if you seek deep and delicious impact, read on.

—**Arran Hamilton**

Kuala Lumpur, Malaysia

—**Douglas Reeves**

Boston, Massachusetts, USA

—**Janet Clinton**

Melbourne, Australia

—**John Hattie**

Melbourne, Australia

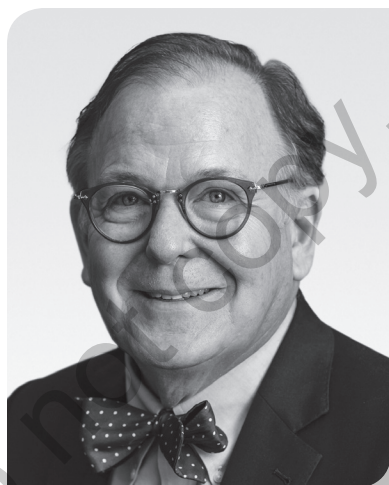
Do not copy, post, or distribute

About the Authors



Dr. Arran Hamilton is Group Director of Education at Cognition Education. Previously, he has held senior positions at Cambridge University Press and Assessment, Education Development Trust, the British Council, Nord Anglia Education, and a research fellowship at Warwick University. His core focus is on translating evidence into impact at scale, and he has overseen the design, delivery, and evaluation of education programs

across the Pacific Islands, East Asia, the Middle East, the United Kingdom, Australia, and New Zealand. Arran's recent publications include *The Gold Papers* and *The Lean Education Manifesto*.



Dr. Douglas B. Reeves is the author of more than 40 books and more than 100 articles on leadership and education. He has twice been named to the Harvard University Distinguished Authors Series and was named the Brock International Laureate for his contributions to education. His career of work in professional learning led to the Contribution to the Field Award from the US National Staff Development Council, now Learning Forward. He was also named the William

Walker Scholar by the Australian Council of Educational Leaders. His recent books include *Deep Change Leadership*, *Achieving Equity and Excellence*, *From Leading to Succeeding*, and *Fearless Schools*. Doug is the founder of Creative Leadership Solutions, with the mission to improve educational opportunities for students throughout the world using creative solutions for leadership, policy, teaching, and learning. Through this he has worked across more than 40 countries.



Dr. Janet M. Clinton is the Deputy Dean of the Melbourne Graduate School of Education (MGSE), Director of the Teacher and Teaching Effectiveness Research Hub, and Professor in Evaluation at MGSE. She has wide international experience as an evaluator and educator and has an extensive publication record. She has led over 120 international evaluation projects across multiple disciplines, in particular education and health. Her major

interest in program evaluation is the development of evaluation theory, mixed methodologies, and data analytics. Janet's current evaluation work focuses on development of evaluation frameworks and implementation protocols, as well as using evaluation as a vehicle for change and building capacity through extensive stakeholder engagement.



Dr. John Hattie is Emeritus Laureate Professor at the Melbourne Graduate School of Education at the University of Melbourne, Co-Director of the Hattie Family Foundation, and Chair of the Board of the Australian Institute for Teaching and School Leadership. His areas of interest are measurement models and their applications to education's problems as well as models of teaching and learning. John has published and

presented over 1,600 papers, supervised 200 thesis students, and published 60 books, including 28 on understanding and applying the Visible Learning research.

CHAPTER 1

D1: Discover



Agree on ONE education challenge that's worth progressing above ALL else

1.1 Establish a Backbone Organization



1.2 Decide the Education Challenge



1.3 Explain the Education Challenge



1.4 Agree on What Better Looks Like



INTRODUCTION

Human history is resplendent with brilliant acts of accomplishment. These include the crafting of the Pyramids, Stonehenge, and the Great Wall of China to more modern feats like carving out the Panama Canal, putting astronauts on the moon, and (more controversially) undertaking the Manhattan Project to unleash nuclear capabilities. We have become deeply skilled at bringing complex goals to life.

None of our greatest feats emerged through osmosis. They each started with a clear goal or challenge. They each leveraged the collective knowledge, capabilities, and technology available at the time. They each involved the mobilization of vast legions of people,

often for long durations. Sometimes for centuries. They were each singular in focus. And none of them got it right the first time. In fact, the path to success was often littered with failure, after failure, after failure. With brilliance then (eventually) emerging from the rubble.

Take the 1969 moon landings. Humanity would never have made it if the goal was a passing whim, if there was no backbone organization to drive learning and implementation (NASA), if the goal wildly exceeded the available technology (i.e., before the invention of rocketry), or if it was one of many competing priorities of the backbone agency (Brinkley, 2019). Imagine what would have happened if alongside the quest to put people on the moon, NASA were also charged with curing cancer, inventing the internet, and building flying cars. Likely nothing, because NASA's success in the 1960s was arguably the result of its single, publicly proclaimed, and unambiguous goal: getting astronauts onto the surface of the moon (and back again) before the end of the decade.

The clear lessons from humanities' greatest past successes are as follows (Kotter, 2012):

1. Have goals that are widely agreed on.
2. Set goals with deep rather than wide ambitions (i.e., trying to push ONE dream rather than multiple dreams).
3. Establish a backbone organization to bring the goals to life.
4. Assume that success won't be achieved easily and possibly not through the means originally intended.
5. Have a rigorous implementation process (e.g., *Building to Impact 5D*).

Before you get moving, you need to ensure that you have identified the right priorities.

If you are reading these words, it is likely because you have an intuition, hunch, or inkling that there is something that could and should be improved in your school or district. Most improvement initiatives start this way. After all, unless you already have a *feeling* that something needs to be done to improve on the status quo, why would you *want* to embark on any form of crusade? However, before you get moving, you need to ensure that you have identified the *right* priorities:

- That the area you have identified is *uniquely* important to improve vs. all the other things you could be doing with your time—including doing nothing at all;
- That it is actually *amenable* to change; and
- That you haven't bitten off *more* than you can ever hope to chew.

In other words, you need to make sure that you have identified the right priority, that it is something that you have the resources and

expertise to meaningfully enhance, and that you are not attempting to pursue too many such agendas simultaneously (i.e., the road to nothing). This is where **Stage D1: Discover** comes into play, which is where you establish your **backbone organization** (1.1.), decide and explain your **education challenge** (1.2 and 1.3), and agree on what better looks like (1.4).

1.1 ESTABLISH A BACKBONE ORGANIZATION

Once you have an *inkling* that you want to accomplish something, you need to establish a backbone organization (Kania & Kramer, 2011) to explore the local landscape to identify and confirm a suitable goal, or what we call the education challenge. There is something a bit chicken and egg about this. On the one hand, you are unlikely to set up a formal structure unless you already have a good idea about what you want to use it for. But on the other hand—if after setting it up, you dive straight into the implementation of plans, programs, and activities—how can you be sure that you are actually spending your time on worthwhile education challenges? The danger is that you end up implementing ill-thought-out initiatives that generate not one iota of meaningful impact.

The backbone team is often a temporary organization. Its sole purpose is to (1) drive the search for appropriate meaty/challenging goals, (2) design initiatives to achieve the goals, and (3) then implement, evaluate, iterate/sustain/scale, or stop. It acts as the central nervous system of your initiative.

If you are working at the district, state, or national level, your backbone team is more likely to comprise full-time/seconded members—who live and breathe the process for 100% of their working hours. However, if you are working at the school or departmental level, it is more likely that your backbone team will be made up of people who are already wearing multiple hats; thus, you are less likely to have the luxury to enable the team to shed all their existing business-as-usual tasks. However, your quest is more likely to be effective if at least one (but ideally more) of your backbone team members has their load significantly lightened so that they can focus their thinking and implementation energies on generating big, deep impact. It also helps if at least one of the members has been explicitly trained as an implementation specialist.

Here are some of the core roles that you will want to have on your backbone team:

1. **Sponsor(s).** The sponsor or sponsorship group is the ultimate owner and endorser of the activity. They are likely to be very senior stakeholders within the organization (e.g., the school leadership team or cluster/district/system leadership), with

the authority and budget to authorize the inquiry, establish the backbone team, and communicate to other parts of the organization that *this is important, you should take it seriously, and you need to engage*. They should also meet with the backbone team regularly to check on progress, provide encouragement, iterate collective thinking, and agree on how they can best use their leverage, authority, and expertise to maintain collective momentum throughout the wider system or the school.

2. **Team leader(s).** The team leader is responsible for leading and coaching the backbone team to identify goals worthy of everyone's time and energy and then to design and implement initiatives with a high probability of generating impact. You might have one team leader or co-leaders. Ideal team leaders can walk the tightrope between being analytic and agile, and it helps if they have a runway of several years remaining with the organization so that they can see things through.
3. **Investigators.** In partnership with the team leader(s), the investigators are going to undertake and/or facilitate the search, design, delivery, and double-back activities. The investigators are the backbone of your backbone organization! They get into schools and classrooms and engage with teachers, students, parents, and the community—gathering data, testing ideas, supporting implementation, and doubling-back on everything.
4. **External facilitator (optional).** An implementation specialist/scientist who has a deep understanding of these processes and wider methodologies is one example of an external facilitator. Their role is to coach and support, to warn you that you are getting into the weeds, and to help you get back out again.

In living organisms, backbones are crucial. In mammals, all the important wiring and information passes through them. But they are part of a wider architecture that co-opts other resources in to do the thinking, moving, lifting, sensing, and feeling. Your backbone team will likely operate in the same way—connecting with and drawing on expertise within and outside your organization to support that thinking, moving, and lifting. And some of your membership might change to reflect the types of expertise needed at different stages of inquiry—adding different types of flesh to the bone. For example, during the Discover (D1) and Design (D2) stages of *Building to Impact 5D*, you will likely need access to more analytic and inquiry-oriented team members to help you search the options in the design space. Think of this as like an (educational) management consulting skill-set. During Stage D3 (Deliver), you also need people with a project management mindset—people who can get things done.

If you are leveraging *Building to Impact 5D* as an inquiry process within the department of a school or more informally within teaching teams, you probably won't need to get as hung up on the

explicit notion of a backbone organization structure and formal role descriptors. You might, for example, already have a well-functioning professional learning community model in place, and you might decide that you want to focus on using the methodology and tools in this book *within* that existing organizational structure. That's perfectly fine.

However, if you are working at the whole school, district, or system level, then explicitly having a backbone organization and giving the members a label and *articles of incorporation* (i.e., license to operate) significantly increases the probability that they and the wider system will take the investigation seriously. These articles outline the purpose of the entity and the rights and duties of the members.

Your articles of incorporation likely include the following:

- A statement of purpose,
- A term limit,
- Membership,
- Roles and responsibilities,
- What it is that people are committing to do,
- How meetings will be organized and chaired,
- How decisions will be made (e.g., voting/consensus, power of veto, etc.),
- Delegated authorities (i.e., what decisions they can make directly and what needs to be agreed on by the sponsors),
- Resources that will be made available, and
- The process for amending the articles of incorporation.

We deliberately provide no template for you to do this because (for your community) the process is likely to feel more considered, important, and legit if it is crafted locally, rather than with a cookie-cutter template for you to fill in the blanks.

You might also go so far as codifying a responsibilities and accountabilities matrix so that everyone is clear what their specific role is within the backbone organization. If your backbone team is relatively small, you may not need to do this. But the larger it becomes, the more likely you are to suffer from the **Ringelmann effect**. This is the tendency for groups to become less effective the bigger they get (Forsyth, 2014; Ringelmann, 1913). There are two aspects to this. First is a loss of motivation (i.e., social loafing) as you assume someone else will do the work, and they do too! Second is a lack of coordination between the various players who all trip over one another. Maximilien Ringelmann first noticed this phenomenon when measuring the amount of effort that participants made during tug-of-war rope-pulling contests. The more players, the less coordination and (even) less pulling.

We illustrate how you could set out a responsibilities and accountabilities matrix for your backbone team in Figure 1.1. Here is a

FIGURE 1.1 • Example Responsibilities and Accountabilities Matrix

CODE	STANDS FOR	THIS PERSON IS
R	Responsible	Responsible for performing the task or creating the product and/or output
A	Accountable	Accountable for and has sign-off authority for the task (e.g., the project manager, sponsor, or technical lead)
S	Supports	Provides expertise, advice, and support to the person responsible for the task or document and others
I	Informed	Informed of task progress or results, usually by the person responsible

TASK	PERSON 1 TEAM LEADER	PERSON 2 INVESTIGATOR	PERSON 3 INVESTIGATOR	PERSON 4 SPONSOR
Developing discovery methodology	R	S	S	A
Leading discovery workshops	I	R	S	A
Developing a long list of potential education challenges	S	R	S	A
Shortlisting education challenges	S	R	S	A
Validating shortlisted challenges	R	S	S	A
Identifying interested stakeholders	I	S	R	A
Gathering insights from interested stakeholders	I	S	R	A
Reporting findings of the discovery team back to interested stakeholders	R	S	S	A

quick rule of thumb: make sure everyone is responsible for at least one thing.

Finally, so that the wider community is aware that the backbone organization has been established, you will also want to think about how its purpose and proposed activities are communicated. Depending on your organizational culture, this might need to include reassurances that the purpose and processes are absolutely not a witch hunt—that the focus is not on deficit theorizing or individual fault-finding, but on unleashing deep, positive impact.

BACKBONE ORGANIZATION DOUBLE-BACK



You also need to explicitly double-back during the activity of establishing your backbone organization. Here are some of the evaluative questions that you can consider:

1. Have we identified the right people for the right roles?
2. Does the team currently have the capabilities to undertake the investigation, or would they benefit from external support?
3. Have we established clear governance and operating structures?
4. Does everyone understand their roles and responsibilities?
5. Do people actually have the time to undertake the actions they have signed up for? How will we make time? (*Hint: Look at the processes in Step 2.4.*)
6. Are we setting up a new bureaucracy that is going to get bogged down in paperwork and busywork or have we got the parameters right for agile and meaningful inquiry that generates impact?
7. Who do we need to communicate the purpose and activities of the backbone organization to? Why are we communicating? What are the messages? Is the plan in place and what happens if we don't implement it?
8. What are the double-back actions that we need to imbed within our backbone organization plans before moving to Step 1.2? Have we incorporated these sufficiently?

1.2 DECIDE THE EDUCATION CHALLENGE

With your backbone team in place, the last thing you want to do is get busy implementing random and ill-thought-out initiatives that

address needs that you may (or may not) have. Instead, you need to take time to carefully explore your environment to identify an agenda that you agree is worthy of concerted and sustained collective action. We call this your *education challenge*. It's the big and meaty goal, crusade, mission, quest, or just cause that you have decided you want to progress above all else.

To avoid the bear trap of doing lots but accomplishing little, your backbone team needs to carefully explore the existing environment to *decide the education challenge* (Step 1.2). So, how do you do this?

There are three key considerations you need to have in mind at the start:

- **Your philosophy** (i.e., whether you are looking for *problems*, *opportunities*, or a current level of adherence to an “ideal standard”)
- **Your values** (i.e., the ethical principles that guide your inquiry)
- **Your methodology** (i.e., when and how you collect data and how you use them)

YOUR PHILOSOPHY

Here are three of the philosophical stances that you can adopt at the very start:

1. **Problem-driven inquiry.** This stance involves looking critically for things that are broken or need fixing, oiling, or upscaling (Pritchett et al., 2013). Here are some examples: “Our school is the worst performing in the district,” “We have too much variability in the quality of impact across all our classes and students,” “Student dropout rates are getting worse,” or “We have a problem with our literacy results”
2. **Opportunity-driven inquiry.** This stance involves being appreciative of all the brilliant things you have already accomplished in your local context, while looking for the next important thing to achieve to make things even greater (Cooperrider et al., 2014). Here are a few examples: “We get good academic results but if we introduce a ‘playful’ learning ethos, we might further increase student engagement and outcomes” or “It’s brilliant that most of our students opt to attend school regularly but maybe we could enhance their achievement levels.”
3. **Standards-driven inquiry.** This involves using explicit success criteria like teaching standards, leadership competencies, and student achievement levels that have been developed for your system as a benchmark to measure against. Here are some examples: “Our lesson observations suggest that our teachers are not uniformly using the state-mandated high-impact

teaching strategies” or “Our kids are—on average—exceeding the state numeracy learning standards.”

When the various Chinese dynasties embarked on the Great Wall project, this was to address a major *problem* (i.e., continuous incursions by heavily armed nomadic bands across long and difficult-to-defend territorial borders; Barfield, 1989). Ditto for the Pyramids, where, by some accounts, the problem was death, to which the “solution” was a resurrection machine that magically launched pharaohs to the spirit world. Of course, this raises wider questions about whether the problem is genuinely fixable and, if it is, whether the proposed solution is the best mechanism. This is why thoroughly testing your thoughts, beliefs, and assumptions prior to implementation is at the beating heart of the *Building to Impact 5D* methodology.

NASA’s quest to get to the moon, by contrast, was arguably *opportunity driven*. There was no specific problem that walking on the moon was directly going to solve. It was “just” a brilliant application of human ingenuity from which all sorts of other technological advances and opportunities unexpectedly emerged, including more than 2,000 NASA spinoff technologies (NASA, 2021). However, an obvious question is whether moon walking really *needed* to be done or whether it might have been better to focus energies on solving the big problems of the era, like high smoking rates and lung cancer deaths, hyperinflation, and rampant poverty. If you decide to adopt an opportunities-driven approach, you will need to be ultra-sensitive to the possibility that you might be pursuing—moon walk-style—brilliant grand plans at the expense of ignoring pressing problems and/or needs that are still on the table.

In contrast, when Olympic divers jump off the high board, their success is judged against *standards-driven* criteria. They don’t receive medals for the speed of their dive but for the technical competency, as assessed by a judging panel who award a score. The high-diving community has established global success criteria for what constitutes “exemplary” execution—from arm and leg position, the amount of twist, and the level of the splash on entry—for each type of “officially recognized” dive (see FINA, 2019). The divers are scored on the degree to which their performance meets (or diverges from) this “ideal.” And the divers themselves watch (and rewatch) videos of their performances to understand the points of divergence and to identify their needed improvements. However, the obvious question is who sets these standards and why should they be taken seriously? They are but an opinion and they are scoring an opinion of an opinion. So if you decide to adopt a standards-driven inquiry, you also need to look at the standards themselves. Here, we have more confidence in the value of student achievement benchmarking, and we have more skepticism (or at least curiosity) about teaching/leadership standards and the degree to which their opinions on “best practice” genuinely correlate with student outcomes.

You start with an opportunities-driven appreciative stance; and you do this because it's much more motivational for everyone to celebrate the power of their prior achievements and to reflect on and be grateful for what they already have.

Given the benefits and challenges of each of these three perspectives, we suggest a *hybrid approach*. That is, you start with an opportunities-driven appreciative stance; and you do this because it's much more motivational for everyone to celebrate the power of their prior achievements and to reflect on and be grateful for what they already have than to immediately be self- and collectively critical. Once you have framed the positives, you can then pivot to a problem-driven inquiry: "We've achieved so much but what issues are still on the table that if we can address them will result in a much better tomorrow?" And concurrently, you can move to a standards-driven approach: "Which standards represent more than just someone's opinion and how are we tracking against them?"

YOUR VALUES

Some of the values that are important to us include the following:

- Privileging evidence over beliefs and opinions
- Deliberately putting our pet beliefs and opinions to the test
- Not misinterpreting the data to make the facts fit our sacred ideas
- Accepting when we are wrong, even when it's emotionally painful
- Respectfully engaging with others
- Being gleeful evaluators of our own impact
- Remembering that it's *all* about growing student learning outcomes
- Embracing the fact that education is fundamentally transformative to lives and life chances. What we do comes with a *major* responsibility.

We think that these values are likely to be equally applicable to you. Most of the above are bound up in the Enlightenment values of reason, rationality, and empiricism that propelled humanity from astrology, flat-Earth theorizing, and human sacrifice to the brilliance of the modern age (Pinker, 2019). But there will also be other values that are especially relevant to you and your local context.

Once agreed on, these values become the test by which all your decisions are made. It's worth you taking some time to tease this out, to include those values in your articles of incorporation, and to regularly swing back to them as you undertake your inquiry. Then when someone in your backbone organization says, "I think we should introduce [**insert shiny program name*]. My friend in the neighboring school swears by it," you can have a collective response like this: "Are we sure we are privileging evidence over beliefs? Let's look at the impact data."

Almost as bad as not searching for data is limiting the search to evidence that supports the assumed hypothesis and deliberately not looking for disconfirming data. This is pseudo-science, or what Richard Feynman (1974) calls “cargo cult science.” Here is another way of putting it: “I think X is true and I will now search for research that supports X. I will not look for anything that disconfirms X and if I accidentally find it, I definitely won’t read it” (see Hattie & Hamilton, 2020a).

YOUR METHODOLOGY

Once you have decided your philosophy and deliberated your values, you then need to consider how you will explore your environment to identify and validate appropriate goals. One non-negotiable thing is that this *always* involves collecting, reviewing, and analyzing data. Without data, you are operating entirely on hunch, instinct, or intuition, and you have no way of knowing whether you have selected an appropriate education challenge or whether you are wasting your energy on low-priority (or even unattainable) initiatives. You’re just another person with an opinion, possibly even a “crazy uncle.”

Here are three of the ways that you can undertake your exploration:

- 1. Ideas-driven approach (a.k.a. the deductive method).** Here, you begin by unpacking your collective *inklings* or *beliefs* about what the important improvement agendas *could* be, and you then collect and review data about each of these hunches to test whether the facts or on-the-ground realities correspond with your beliefs. For example, “We think that migrant children in our school are being short-changed. Is this the case?” or “We want to make learning more fun. Is it true that learning isn’t fun for our kids, and does it actually need to be? Or is ‘real’ learning generally painful and challenging?”
- 2. Data-driven approach (a.k.a. the inductive method).** Here, you keep your preconceived ideas firmly locked away and instead collect and analyze all the data you can get your hands on and search for patterns, needs, and opportunities in the data. Some examples are as follows: “The data are telling us that student achievement in standardized assessments has declined over the last 5 years. Should we be concerned about this?” or “The data are telling us that our teacher retention rates have increased significantly. Should we be pleased?”
- 3. Reason-driven approach (a.k.a. the abductive for parallelism method).** This starts from your seeing some problem, surprising details, or some event that cannot be readily explained. Abduction invites the educator to choose the “best” explanation among alternatives to explain this

conundrum. And if it does not convince, then move to the next “best” explanation. Here’s an example:

I note the brightest students seem to be cruising. My explanation is that they need gifted classes. But upon trying them they are still cruising, so I ask teachers in regular classes about their concept and implementation of “challenge.” I find their view is that challenge is coverage more of the curricula with these students, but the students want more challenge in the depth and not merely move coverage at a surface level.

During Step 1.2, it is not *yet* necessary for you to use the reason-driven approach to explain the root causes of each potential education challenge. You may have identified quite a long list of potential problems and/or goals. It is more efficient to wait and then to apply abductive reasoning *only* to the challenge(s) you have subsequently decided to prioritize. The whole of Step 1.3 leverages this abductive method—because in order to improve, we need to understand what we are building on top of.

If you have the resources, you could undertake your exploration in all three ways. First, have a mini-search party that starts with their hunches, and looks for confirming and/or disconfirming data. Then another search party tries their best to keep their hunches firmly locked away and instead looks for the signal in the (data) noise. Finally, develop another list of possible explanations to then rule out if the evidence does not support them (but again, it is not essential that you build explanations for the identified education challenges until Step 1.3). You can then compare what each party comes up with and see whether they vector in on common agendas.

Whichever way you choose, some of the sources of data that you will want to explore include those identified in Figure 1.2.

Depending on your national context, there may also be data available from external reviews or school inspections.

THE MANY, THE FEW, THE ONE

In our work, we have had the privilege of working directly with and learning from thousands of schools and more than 50 system-level agencies. Most of the schools and systems we have collaborated with were skilled at analyzing their present situation and identifying (very) long lists of potential goals and problems as well as improvements and initiatives. However, what many were less good at was whittling these down to an achievable number of priorities.

FIGURE 1.2 • Sources of Data

	NATIONAL/ STANDARDIZED ASSESSMENTS	INTERNAL ASSESSMENTS	LESSON OBSERVATIONS	INTERVIEWS AND SURVEYS	ELECTRONIC MANAGEMENT INFORMATION SYSTEM (EMIS) DATA
Benefits	<ul style="list-style-type: none"> Usually reliable; centrally/professionally designed, assessed, and moderated Provides comparative data across years and decades No extra effort required to collect 	<ul style="list-style-type: none"> Usually more valid/connected to the local learning context More amenable to item-level analysis (i.e., can unpack the areas where students performed well or need additional support) 	<ul style="list-style-type: none"> Direct insights into classroom activity Also has developmental value as part of professional inquiry 	<ul style="list-style-type: none"> Ability to gather perceptions from students, teachers, and/or community members Can tailor to specific inquiry areas 	<ul style="list-style-type: none"> Provides wider data on topics such as these: <ul style="list-style-type: none"> Student attendance Teacher retention Teacher sickness Health and safety Financial management Student characteristics and needs
Challenges	<ul style="list-style-type: none"> Only available for exam years Often reports holistically (i.e., cannot always access item/question-level data) 	<ul style="list-style-type: none"> May not be as reliable as national assessments Extra time/effort/cost to design, administer, and assess May not provide comparable data across years 	<ul style="list-style-type: none"> May not be representative of day-to-day activity in the classroom Presence of an observer may influence behavior of teacher and learners Difficult to “see” student learning with your eyes/open to different interpretations by different observers 	<ul style="list-style-type: none"> Extra workload to design, administer, and interpret Willingness of stakeholders to engage in the process Provides insight into perceptions and/or beliefs rather than impact and outcomes 	<ul style="list-style-type: none"> Often difficult to benchmark your school’s data against other schools, as these data are often not in the public domain
Best for	Big-picture longitudinal analysis	Vector analysis of specific cohorts, classes, curriculum areas, and topics	Holistic understanding of what is happening in individual classes	Understanding stakeholder perceptions and for gathering ideas	Understanding input drivers that might impact teaching and learning in classrooms

Source: Adapted from Sharples et al. (2018).

They often really struggled to take things off the table and wanted to progress all of their identified education challenges. But being highly selective is extremely important. In our own school and system improvement work, we have consistently found that the smaller the number of priorities, the higher the probability of impact. To repeat, *the smaller the number of priorities, the higher the probability of impact*. You need to be extremely ambitious about a very small number of things, or in the words of Viviane Robinson (2018) in her book of the same name: “Reduce change to increase improvement.”

Previous research based on an analysis of more than 2,000 school plans, along with 3 years of student achievement data for those schools, suggests that the optimal number of initiatives is no more than six (Reeves, 2013). Some of the world’s premier organizational consultations from Booz Allen and PwC (Leinwand & Mainardi, 2011) came to similar conclusions. In a survey of more than 1,800 organizations, the researchers found that the majority said that they had too many initiatives, including some that were conflicting, and more than 80% said that initiative overload was wasting time. They concluded that the optimal number of initiatives was three to six. That represents a confluence of evidence between the worlds of education and other organizations. Thus, when we see the common occurrence of dozens of priorities in school and district improvement plans, we are dismayed at yet another example of how prevailing and consistent evidence in favor of focus is ignored.

Where schools and systems have long lists of priorities for improvement, the less likely it is that stakeholders

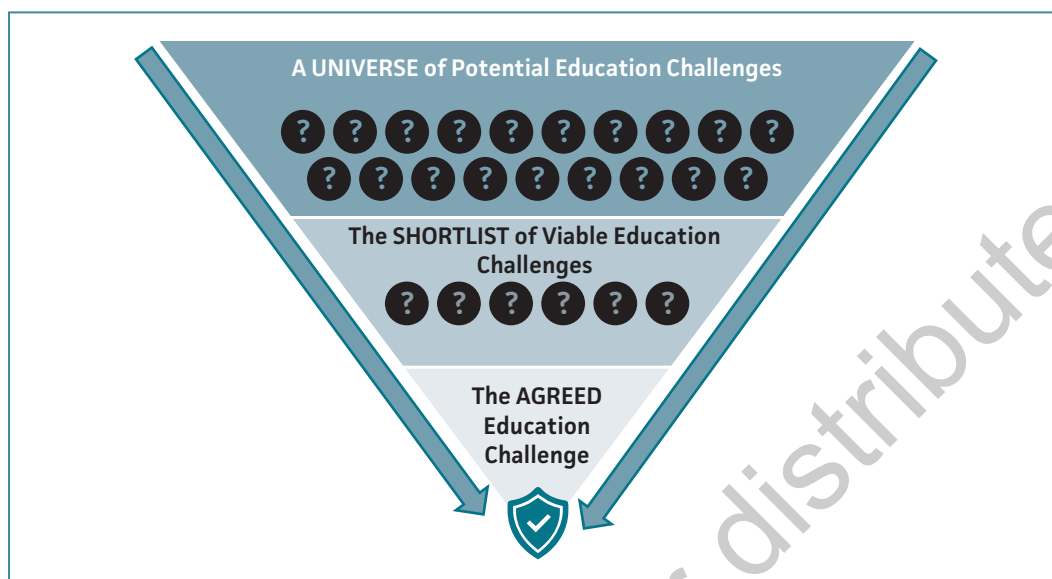
- understand the big picture, let alone agree with it;
- understand how the various initiatives that are being implemented contribute to making things better within the frame of that big picture;
- can carve out time to make meaningful progress in leading or contributing to these multiple agendas;
- can systematically measure and reflect on their progress; and
- can identify where to next.

You need to focus on less to achieve more. This means that you need to find a way to progress from a universe of near-infinite education challenges or goals that all seem worthy of your time and energy to a shortlist of viable challenges and to then boil this down to ONE agreed priority. In other words, from the many, to the few, to the ONE. We illustrate this in Figure 1.3.

You need to focus on less to achieve more.

If you are thinking that one priority is too restrictive, you’re right—it’s deliberately so. For you to be able to agree on one thing, it normally ends up being a big all-encompassing sort of challenge that

FIGURE 1.3 • The Education Challenge Funnel



Note: Copyright © Cognition Education (2022). All rights reserved.

requires the design and implementation of several work strands, initiatives, or programs to tackle different aspects of the improvement area. So while Doug Reeves originally advocated the *rule of six* in some of his earlier writings, we now collectively propose the *rule that one priority will probably end up sprouting into six contributing initiatives*. So, it's better to start with one priority because 1 education challenge \times 6 initiatives is more realistic to manage than 6 education challenges \times 6 initiatives.

But how do you home in on the ONE? For every education challenge you identify, you almost need to be looking for reasons not to do anything at all. You might find it useful to ask questions like these:

- What's the absolute worst that could happen if we did absolutely nothing?
- What's the likelihood this worse-case outcome will actually happen?
- Does it really matter if it does happen and/or continues to happen?
- Why should we or anyone care?

We completely get that it seems shockingly cruel to try and conjure up reasons not to address burning platforms or not to embrace big audacious goals. But there is method in our madness. We just want you to select among your competing priorities with

great care—to increase the probability that the one you progress actually matters most and ultimately gets done.

The idea is that you whittle and winnow down until you get to the (hopefully) small handful that stakeholders are extremely reluctant to take off the table. When you ask, for example, whether it really matters whether one-third of children are leaving your school or school system without functional literacy and the collective reaction is complete horror, anger, or incredulity, then you just might have identified a potential education challenge that sufficiently rouses sentiments for all stakeholders to be prepared to make great changes and personal sacrifices to act on.

In Figure 1.4 we share a worked example of this potentially heated and cognitively jarring priority review process.

One of the benefits of this approach is that it encourages you to take a long-term perspective and think beyond test grades to long-term life chances. We know from research on the **Matthew effect** that children not achieving a year's learning growth for a year's teaching input often spiral into a doom loop that is more likely—in adulthood—to translate into higher periods of unemployment, higher incidences of clinical depression and substance abuse, higher incidences of incarceration, and lower life expectancy (Hamilton & Hattie, 2022). So, it is important to keep sight of the fact that today's low-test scores are not a "1-year problem" but a compounding multidecade problem. *That* vision, as awful as it is, should be front of mind as you decide which education challenges are worth your collective effort.

If you are still struggling to vector in on the ONE, it might help you to undertake head-to-head matched pairs comparisons. We illustrate this approach in Figure 1.5.

This process involves your backbone team and potentially also wider stakeholders coming together to review each of the competing priorities to decide which is more important. The idea is that the process is repeated until every potential focus area is compared against every other. You might do this comparison through group discussion and by coming to a group consensus, or you may choose to literally vote individually on each priority and tally up the collective wisdom of the crowd.

One of the interesting things that often emerges from this process is that stakeholders usually begin to express thoughts like, "Maybe the shortage of qualified teachers *could* be one of the reasons why such a high proportion of our kids are functionally illiterate?" These discussions about whether some potential challenges could actually be *causes* and others *symptoms* or *effects* are really important—in terms of both deciding what the real priority should be and also in understanding what factors could be contributing to the existence of any identified problems.

FIGURE 1.4 ● Education Challenge Hunting 

POTENTIAL EDUCATION CHALLENGES	DOES IT REALLY MATTER? (I.E., THE SO WHAT? PROVOCATION)	WHAT'S THE WORST THAT COULD HAPPEN IF WE DID NOTHING? (I.E., WHERE IS THE EVIDENCE? HOW STRONG IS THE EVIDENCE?)	SHOULD WE CARE? (I.E., IS THIS OUR MOST IMPORTANT MEGA-CHALLENGE?)
Our school buildings are shabby and overcrowded	<i>Is there a strong relationship between the quality of infrastructure and class size and student attendance/achievement?</i>	No statistically significant difference in learning outcomes or attendance between “shiny” and “shabby” schools in Visible Learning Meta ^x . Reducing class size also seems to have only a modest impact on student achievement but at relatively high cost.	Not right now. But we might need to explain to community stakeholders why they should worry less, too.
Difficulty in recruiting qualified mathematics and science teachers—76% do not have an advanced qualification in the subject they teach	<i>Do teachers really need to have an advanced qualification in the subject they teach?</i>	Mathematics and science teachers with an advanced qualification are an average effect size $d = 0.10$ more effective.	No. This is probably not big enough to warrant the investment right now.
33% of our 13-year-olds are functionally illiterate	<i>Is it essential that everyone is literate? Can we live with 66%?</i>	Illiterate adults in our context are <ul style="list-style-type: none"> • 4 times more likely to be in low-paying jobs • 2.7 times more likely to be convicted of a criminal offense • 3 times more likely to die before age 60 	This feels pretty bad. But maybe they will catch up?

FIGURE 1.5 ● Matched Pairs Comparison

Difficulty in recruiting qualified mathematics and science teachers	vs.	33% of our 13-year-olds are functionally illiterate
---	------------	---


Another potential approach to education challenge identification is what we call *So What and So Prove It*, which we illustrate in Figure 1.6.

As you undertake your whittling and winnowing, you are also going to want to gather and rank the quality of the evidence that you are using to help you make decisions. Do you have

- **High confidence** that the backbone team has unearthed a significant trove of quantitative and qualitative data from local and global sources, which all pull in the same direction;
- **Medium confidence** as above but with some gaps, or that data seem to pull in different directions; or
- **Low confidence** because judgments are being made based on anecdotes, hearsay, and intuition?

For example, if you look carefully at Figure 1.6, you will see that the quality evidence for the *So Prove It* statements varies considerably. Many are just assertions or opinions without underlying data in their support.

If at the end of your whittling and winnowing, you cannot identify a single priority that the majority of stakeholders strongly agree is worth collective action (and collective pain) to resolve, then you might want to consider disbanding your backbone team. You are very unlikely to make progress on goals that are either non-systemic or where sentiments are not sufficiently aroused to drive and sustain the hard graft of implementation.

FIGURE 1.6 • So What and So Prove It 

EDUCATION CHALLENGE: 33% OF OUR 13-YEAR-OLDS ARE FUNCTIONALLY ILLITERATE	
So What 1: They can't read!	So Prove It 1: Internal assessment data demonstrate that 33% of our students are not achieving the expected literacy standards.
So What 2: They will fall further behind.	So Prove It 2: If they can't read, they can't engage with the curriculum sufficiently.
So What 3: They are more likely to drop out of school.	So Prove It 3: We ran the numbers across the district. Children who did not meet the functional literacy standards are five times more likely to drop out.
So What 4: They are less likely to get good jobs.	So Prove It 4: National employment data correlate illiteracy with a significantly higher probability of unemployment and lower lifetime earnings.
So What 5: They are more likely to have a lower life expectancy.	So Prove It 5: Correlational data show an association between lower literacy levels and earlier onset of comorbidities.

Source: Copyright © Cognition Education. (2022). All Rights Reserved.

You also want to think deeply about whether your selected challenge is genuinely amenable to change. For example, if NASA had been established in 1860, it would have made much less progress with its Moonshot goals and would likely have quickly given up because the supporting rocketry and computer technology were a long way from being developed sufficiently.

However, if you have identified a worthy and amenable challenge, the next key step is to define it and explain the causal elements. This is also known as *building a theory of the present*, because you cannot improve or enhance the future without properly understanding the present from which it builds.

1.3 EXPLAIN THE EDUCATION CHALLENGE

Imagine waking up in the night with searing back pain and hobbling to the doctor the next day. It is extremely unlikely that your chosen physician will pull out a random treatment option without a diagnosis. Instead, they have been trained to do two key things before even *thinking* about potential options (Wilson, 2012):

1. **Understand and define your problem or need as clearly as possible.** The doctor might ask questions like these: What is the severity of the back pain? How much does it interfere with your daily life? Does it affect the whole back or just a specific part? Which part? Are there certain times of day, temperature conditions, or types of activities that are more painful? Is it muscle or joint pain? Has it gotten progressively worse over time? Have you ever had it before? Does it run in the family?
2. **Build a hypothesis about the causes.** Painful symptoms do not usually emerge on their own. They generally present at the end of a causal chain that starts in some other bodily system and then ricochets through different cells and organs to manifest as “back pain.” The cause *might* be arthritis, a kidney infection, bad posture, pregnancy, a sports injury, or tens of other possibilities. Your doctor might order X-rays or imaging scans to find out more. But even if the scans identify, say, arthritis, the causal chain does not stop here. Arthritis *might* have been caused by environmental factors, like a lifetime of picking up boxes resulting in joint damage; by a genetically primed immune response; by being overweight; or by many, many other influences. Without having a plausible working model of the causes, your doctor cannot prescribe the right sorts of treatment.

This type of medical reasoning is called **abduction**, which is the third sister to **induction** and **deduction** (Hattie & Larsen, 2020). The same abductive principles apply to education improvement. If we

have not defined the target area (i.e., the education challenge) and if we have not built a working model of the causes, we are flying blind. We would be like doctors randomly prescribing muscle relaxants when a more careful diagnosis might show that what's really needed is weight loss and dietary change, or even surgery.

Given that whatever changes we make in schools and systems incur an opportunity cost (time, money, buy-in, good will, etc.), we need to be as sure as we can be that the treatments, interventions, programs, activities, or designs that we propose will actually push *our* needle—as opposed to pushing needles on agendas we do not currently have. This means that we need to employ the same abductive diagnostic protocols as good doctors.

DEFINE THE EDUCATION CHALLENGE

Let's assume that during your whittling and winnowing process, you selected 33% of our 13-year-olds are functionally illiterate as your education challenge. This is a bit like going to the doctor and saying you have a nondescript type of back pain. So, like the doctor, you need to interrogate and expand your definition. Here are some of the questions you might find it useful to ask:

1. What is our **definition** of the challenge?
2. What **data** have we used to inform our definition?
3. How **valid and reliable** are the data?
4. What is the **span** of our challenge? That is, does it relate to all, many, some, or a few stakeholder groups?
5. What are the **stakeholder characteristics** of those that the challenge most relates to (e.g., males or females, younger or older, teachers or students, resource rich or resource poor, etc.)?
6. What is the **severity of impact** on the different stakeholder groups?

We call this the **breakdown structure**, and we illustrate what the outcomes of this process might look like in Figure 1.7.

What do you notice as you scan Figure 1.7? Take time to reflect and then compare your thoughts with our observations as follows.

First, it's clear that the challenge *might* be under-recorded, because the measuring tool *could* be letting some false positives slip through the net. Equally, it might also be over-reported because the school currently assumes that children who have not taken the literacy assessment would also have failed it. Both these observations suggest the need for more digging and possibly also for more robust student assessment and reporting mechanisms to be put in place.

FIGURE 1.7 • Education Challenge Breakdown Structure

EDUCATION CHALLENGE	
33% OF OUR 13-YEAR-OLDS ARE FUNCTIONALLY ILLITERATE	
DEFINING CHARACTERISTICS	COMMENTARY
1. <i>Functionally literate</i> is defined as achieving 75%+ on our school's internal literacy assessment. Functional <i>illiteracy</i> = scoring less than 75%	External consultants have recently benchmarked our internal assessment and suggested it is of low validity and reliability. The challenge could actually be much bigger (!!!).
2. In our school, 15% of 13-year-olds not meeting the standard have not actually taken the assessment.	Can we assume that all those who have not taken the assessment would have failed to achieve the 75% threshold?
3. Students who have not taken the assessment are <ul style="list-style-type: none"> ○ Disproportionately boys (68%) ○ Disproportionately in rural areas (74%) ○ Disproportionately from low socioeconomic status (SES) backgrounds (83%) 	Is this a consequence of teaching quality in our rural districts or of lower levels of community engagement and/or aspiration?
4. Of the students who took the assessment and did not achieve the 75% threshold, they were: <ul style="list-style-type: none"> ○ Disproportionately boys (57%) ○ Disproportionately from low SES backgrounds (83%) 	Can we find any data on teacher demographics? For example, are these students more likely to be taught by less experienced teachers or are their teachers more frequently absent?
5. There is a wide dispersal in the degree to which students missed the 75% threshold: <ul style="list-style-type: none"> ○ 32% missed it by 5% or less ○ 41% missed it by 5–15% ○ 27% missed it by more than 25% 	Putting to one side that our assessments may not be as valid or reliable as external assessments, the data suggest that there are varying degrees of student support required. For example, for 32% of students, it might be the case that they only require light-touch additional support.

Second, there do seem to be some patterns among the students who have not met the assessment threshold. It seems to affect boys and children from lower socioeconomic status (SES) backgrounds, so we need to start building a theory about why this is. Do all these children share the same teacher? If so, it could be a characteristic of the teaching. If they have different teachers, it could be related to out-of-school factors. Or it could be that while the teaching approach connects with other learners, there is something about it that's off-putting for these particular learners.

Third, there is variability in how affected students are: 32% of the children who did not meet the literacy standards were actually

within a hair's breadth. They might only require a little additional support and guidance to get them over the line. Yet the type of support required for the 27% of students who missed the target by 25% or more is likely to be more intensive.

So, we already have three lines of inquiry: (1) review the testing instruments and reporting assumptions, (2) consider whether and why children from specific backgrounds are more likely to have support needs, and (3) consider whether the package of support can be varied depending on the severity of need.

BUILD A HYPOTHESIS ABOUT THE CAUSES (A THEORY OF THE PRESENT)

By defining your education challenge with clarity, you have already taken the first step toward building a **theory of the present** and thinking about viable interventions. But we recommend you go deeper. By exploring intensely, you will better know whether your “backache” is caused by bad posture, arthritis, or a bone tumor. You can then select the right treatment pathway (e.g., posture correction, anti-inflammatory medicines, or radiotherapy). And if it is caused by multiple things, you can then mix and match your intervention options.

In addition, there are often multiple layers even within a single causal chain. For example, (1) engrained habits lead to (2) slouching, that leads to (3) a slipped disk, that causes (4) a pinched nerve, and results in (5) excruciating pain. And it may be that the selected interventions tackle many of these links in the causal web, such as paracetamol for the pain, physical manipulation or even surgery for the slipped disk, and cognitive behavioral therapy to adjust your department.

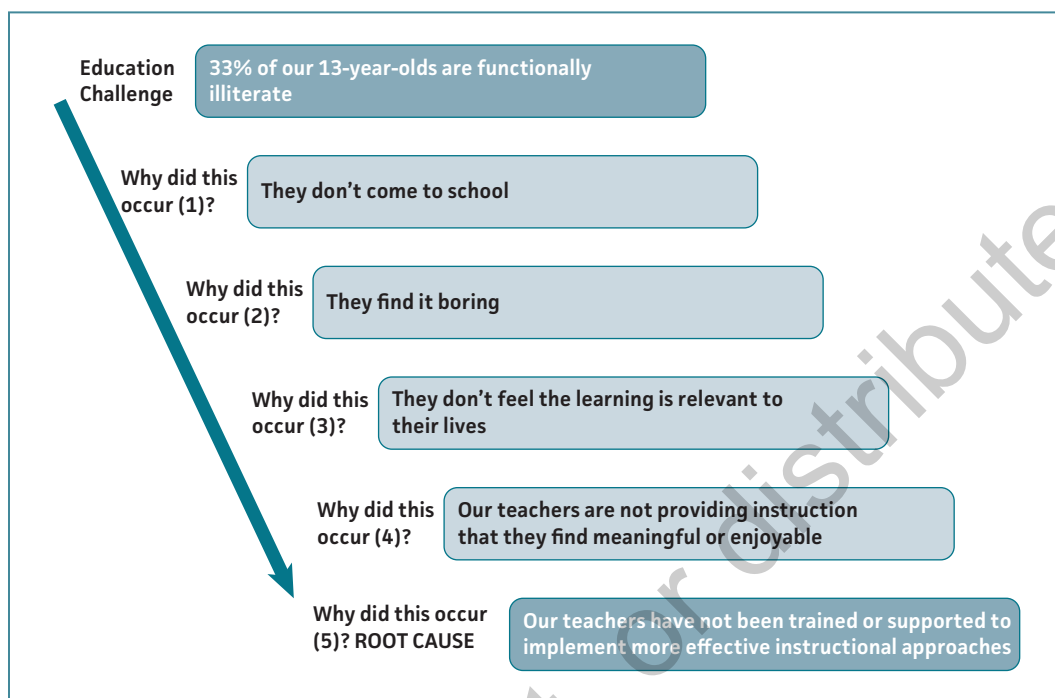
The exact same causal thinking processes can be applied to education improvement. By moving beyond folk theorizing and toward building a robust understanding of our present, we are far more likely to be able to reach for the right interventions that support improvement across the whole causal chain.

One excellent tool for doing this is called the *Five Whys*. It was developed by Sakichi Toyoda and used within the Toyota Motor Corporation (Ohno, 1988). It involves continually asking “Why X is the case?” We provide an illustration in Figure 1.8.

The idea is that you ask these *why* questions over and over and that you use the outputs to build a causal theory. Depending on your inquiry topic, your team composition, and stakeholder beliefs, it is extremely likely you will build many different *Five Whys*, each telling a different story or pointing to a different causal chain. Some of these narratives might be about children's home backgrounds. Others might be about the language or

By moving beyond folk theorizing and toward building a robust understanding of our present, we are far more likely to be able to reach for the right interventions that support improvement across the whole causal chain.

FIGURE 1.8 • The *Five Whys* Approach



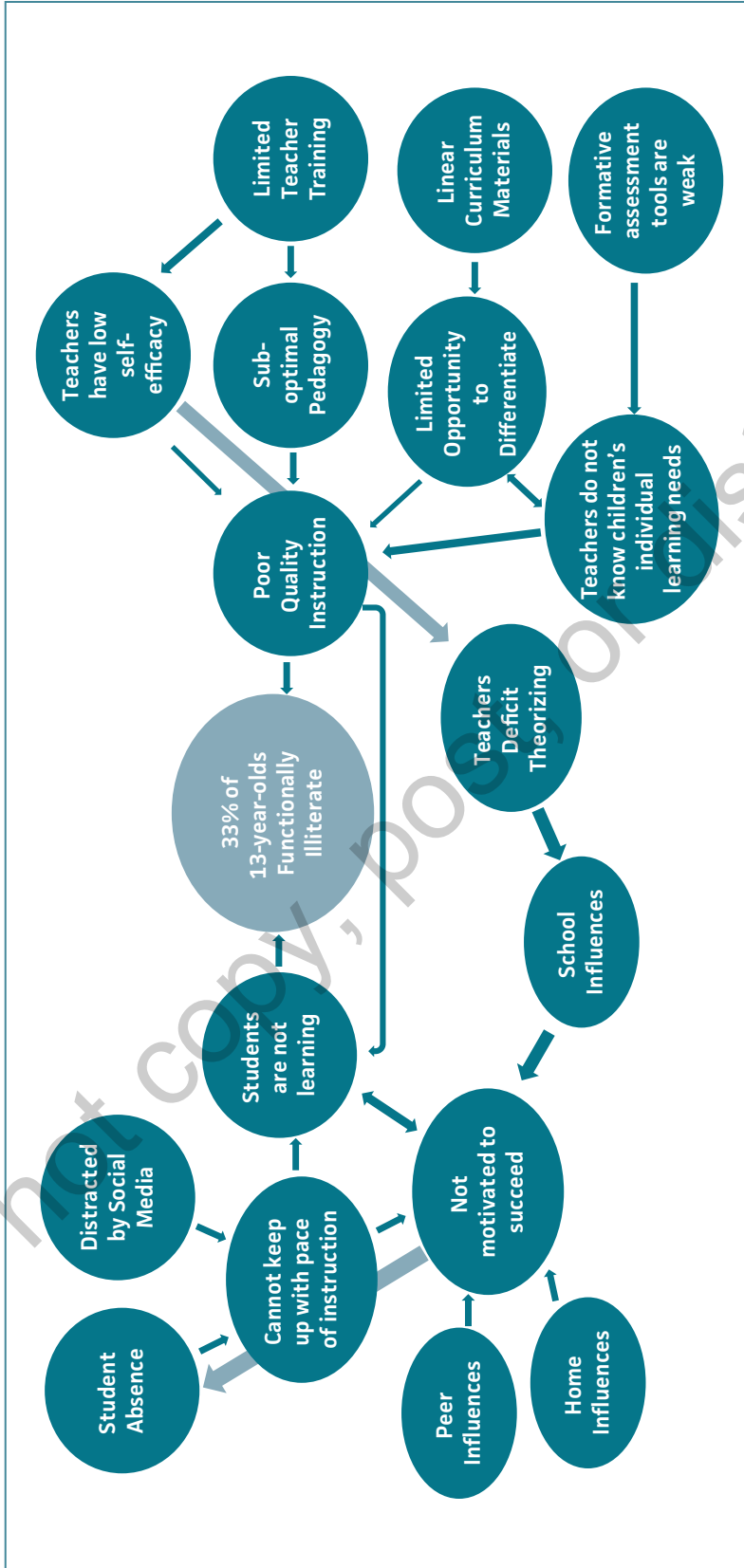
Source: Copyright © Cognition Education. (2022). All rights reserved.

medium of instruction, yet others about instructional approaches (i.e., pedagogy), and others still about classroom relationships, trust, empathy, and well-being. Some might also be about short-term variables like COVID-19 school closures, where the original issue has (hopefully) ceased to exist but where there could still be long-lingering legacy effects.

In Figure 1.9 we illustrate how these different stories can be woven together into a **path analysis** (i.e., theory of the present). This is not a full map, but it gives you an idea of how you might start to blend your many *Five Whys* into a deep and rich explanation of your present. Note that this example is presented from the perspective of an external consultant. If you undertake this type of mapping internally, you may wish to be considerably more appreciative (i.e., sugar coated) in how you label the *influence bubbles*—particularly when you engage with target improvement groups. No teacher wants to hear that they employ “poor-quality instructional approaches” or that they “deficit theorize.” Talking in such language is the pathway to defensiveness, mutual blaming, and people barricading themselves in their minds and classrooms.

One of the benefits of path analysis is its freeform nature. A second benefit is that you can easily build it with sticky notes on

FIGURE 1.9 • Path Analysis



Source: Copyright © Cognition Education (2022). All Rights Reserved.

a whiteboard and draw the connector arrows with a marker pen. However, if you need more structure, other tools you can leverage include *Fishbone* and *Issues Tree* mapping (see Figure 1.10).

However, a disadvantage of both the *Fishbone* and *Issues Trees* approaches is that they tend to encourage linear thinking. The arrows only move in one direction—from cause to effect. However, some causes might also impact other causes (this is known as mediating and moderating in the technical jargon), and some might have greater influence or power of effect than others. This is why we prefer path analysis, which was first developed by Sewall Wright (1921) and is vigorously championed today by Judea Pearl and Dana Mackenzie (2020). You can move the arrows in any direction and even have them going to multiple places. And you can also vary the size of your “bubbles” to reflect your hypothesis about how important each is in the overall scheme of impact.

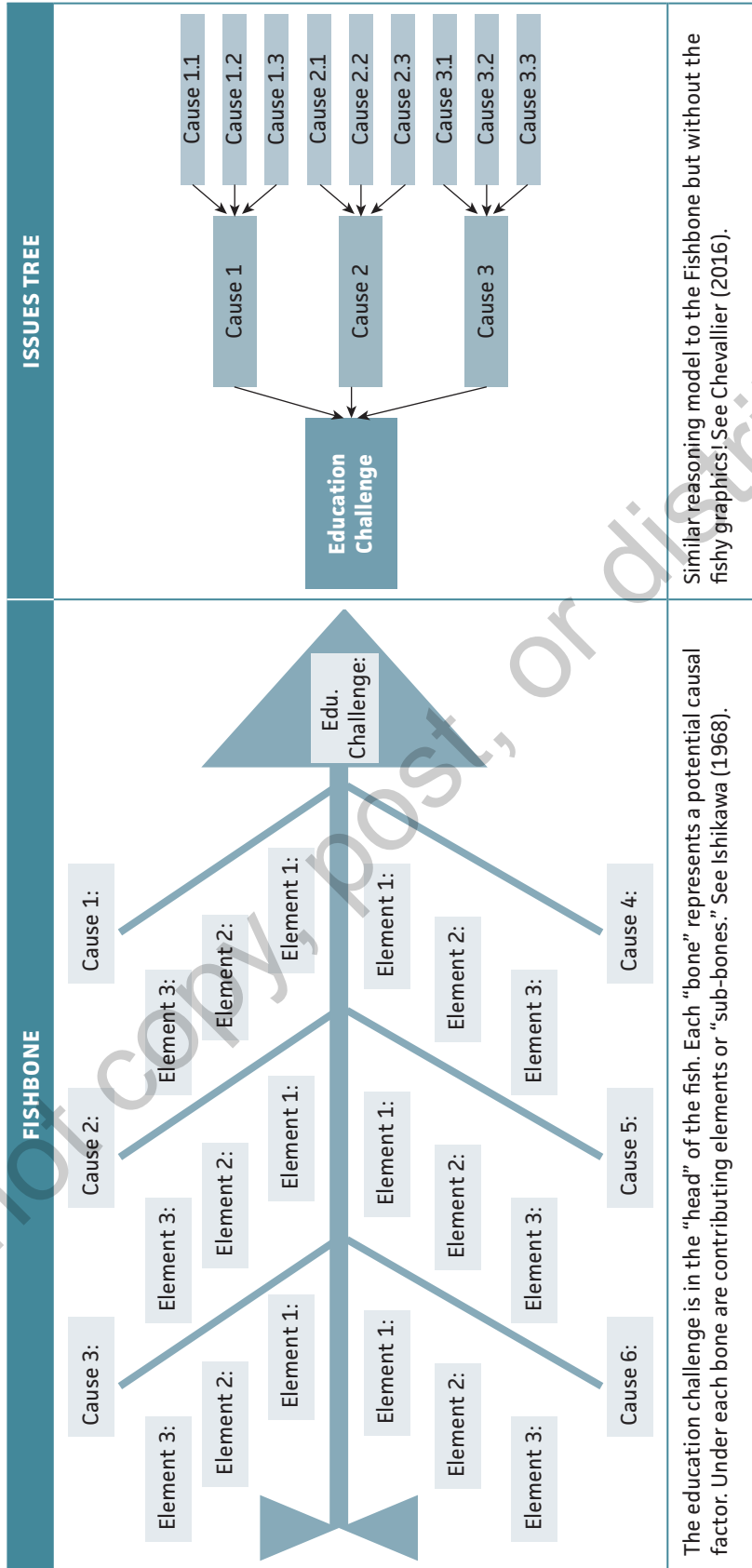
This process is also a form of *abductive reasoning* and is a lot like the way a police detective undertakes a criminal investigation. Often the detectives wheel out a large whiteboard: at the center is the crime or the victim. In orbit around the “vic” are an array of potential suspects (or “perps”) and their respective means, motives, and opportunities. Each of these chains tells a different story. Maybe the crime was perpetrated by Colonel Mustard in the library with the dagger because of a longstanding grudge? Or by Professor Plum in the ballroom with the candlestick because his scandal was about to become exposed?

The detective builds a range of such theories and then searches for evidence that either confirms the suspect as a person of interest or eliminates them from the inquiry. That search for evidence involves the collection of witness statements, review of surveillance video, confirmation of alibis, and so on. When Professor Plum presents flight tickets showing he was out of town the day of the crime and this is corroborated by airport camera footage, his picture is removed from the whiteboard.

During this stage of the inquiry, you are *exactly* like an educational data detective. With your “crime” (or education challenge) at the center of the board and your hunches about different causal chains, you are gathering data to validate (or dismiss) your theory of the present. You are *doubling-back* to test the evidence chain. However, one thing makes this a little more complex than the world of the police detective: we can rarely explain an “educational crime” with a single causal chain such as “Mustard, library, and dagger.” More often, there is a range of interconnected influences—each applying force to varying degrees and contributing to the outcome.

We get that, yes, you may be rolling your eyes at the prospect of undertaking this data detection and that instead you might be itching to just get on with implementing something, anything—because

FIGURE 1.10 • Fishbone and Issues Tree Mapping



Source: Copyright © Cognition Education. (2022). All rights reserved.

you think the causes of your “education crime” are, well, obvious. But if you don’t have a robust (and tested) theory of the present, you are likely to select completely the wrong activities or interventions. You will not then push the needle one iota and you may as well have done nothing. Wild-goose chase galore.

In Figure 1.11, we illustrate how you can collate and review each variable in your path analysis, leveraging a range of data including eMIS, learning walks, student voice, lesson observations, and surveys. This is the built-in double-back.

The idea is that as you explore the supporting data for each link in your causal map, you

- remove influences that have *not* been verified,
- add *new* influences that emerge from your action research, and
- adjust the position of the arrows as you reconsider which variables push energy and in which direction.

Your finalized map is *crucial* to the design activity that you will undertake in Stage D2 (Design), where you will be crafting activities and initiatives that are explicitly designed to interact with and impact the key causal variables. This means that you can then block, weaken, and/or ideally reverse their impact.

If you are working at an individual school level, a single path analysis should do it. However, if you are operating at the district level

FIGURE 1.11 • Validating the Causal Drivers 

CAUSAL VARIABLE	CAUSAL HYPOTHESIS	SOURCES OF VERIFICATION DATA	OUTCOME OF VERIFICATION	VARIABLE REMAINS?
Student absence	Because they are missing from class, they fall behind and cannot catch up	<ul style="list-style-type: none"> • Student attendance data 	Not verified. The level of student absence for our target group is no different from the high-performing students.	Removed
Cannot keep up with pass of instruction	Students do not have sufficient prior knowledge to scaffold and bridge to new curriculum content	<ul style="list-style-type: none"> • Formative assessment data • Student voice 	Verified. Both the assessment data and student interviews suggest 70%+ of the target group does not have appropriate prior knowledge.	Remains

(Continued)

FIGURE 1.11 • (Continued)

CAUSAL VARIABLE	CAUSAL HYPOTHESIS	SOURCES OF VERIFICATION DATA	OUTCOME OF VERIFICATION	VARIABLE REMAINS?
Poor-quality instruction	Students are not making progress because they do not enjoy their classes and do not believe what they study is relevant	<ul style="list-style-type: none"> • Student voice • Lesson observations • Curriculum review 	Verified. Students consistently reported that they found lessons unengaging. Lesson observations also suggested disengagement.	Remains
Limited opportunity to differentiate	Teaching is linear and does not cater to differences in children’s prior knowledge	<ul style="list-style-type: none"> • Lesson observations • Curriculum review • Student voice 	Verified. Limited use of grouping/ differentiation strategies or assessment data to inform teaching.	Remains

or higher, you may find you need more than one map, reflecting the fact that similar symptoms across your schools might have quite different root causes.

1.4 AGREE ON WHAT BETTER LOOKS LIKE

The good news is that we are now (almost) at the end of the Discover Stage (D1). You have identified an education challenge worth progressing and now also have a (pretty) good idea about why it exists. The final thing we recommend you do before moving to the Design Stage (D2) is to start thinking about what success looks like.

By visualizing and articulating what success means in your context, you are more likely to stay focused and motivated to bring it to life. The further piece of good news is that (at this stage) you don’t need to go overboard on identifying measures and setting targets. You will swing back to this bit in Stage D2. However, you can discuss and provisionally agree on the following:

- **An empowering aspirational statement** (e.g., *ALL children at this school will attend regularly, will enjoy their classes, and will achieve the district literacy standards*)
- **Provisional metrics** (like those listed in Figure 1.12)

Things that get visualized remain important and get done!

FIGURE 1.12 • Establishing Provisional Success Criteria

CURRENT SITUATION	“TO BE” SITUATION	BY WHEN	MEASURED HOW
33% of our 13-year-olds are functionally illiterate	88% of 13-year-olds are functionally literate	December 2025	Percentage of total school-age population that achieves Level 3 threshold in national literacy assessment
JUSTIFICATION OF SELECTION OF “TO BE” VALUES			
A review of regional comparator data suggests that an average of 90% of students achieve functional literacy in those other contexts.			


D1: Discover Summary

You have now reached the end of the Stage D1: Discover processes. During this stage of your inquiry:


You will have agreed on ONE education challenge that’s worth progressing above ALL else

You will have done this by:


1.1 Establishing a Backbone Organization




1.2 Deciding the Education Challenge



1.3 Explaining the Education Challenge



1.4 Agreeing on what better looks like



In the next chapter we shift our focus to D2: Design.

Do not copy, post, or distribute