



LEADERSHIP
[for]
TEACHER
LEARNING

*Creating a Culture Where All Teachers
Improve so That All Students Succeed*

DYLAN WILIAM



Table of Contents

Acknowledgements	v
About the Author	vii
Introduction	1
<i>Chapter 1</i>	
Why We Need to Raise Student Achievement, What's Been Tried and Why It Hasn't Worked	7
Why We Need to Raise Achievement	7
The Changing Nature of Work	8
What's Been Tried	20
The Danger of Trying to Copy Others	21
<i>Chapter 2</i>	
Teacher Quality: Why It Matters, What It Is and How to Get More of It	29
Teacher Quality Versus Teaching Quality	30
Improving Teacher Quality	35
Conclusion	62
<i>Chapter 3</i>	
Learning From Research	63
Lesson Study	66
Educational Neuroscience	69
Systematic Reviews of Research	71
Conclusion	98

Chapter 4

Formative Assessment	99
The Origins of Formative Assessment	101
The Definition of Formative Assessment	103
Formative Assessment and Self-Regulated Learning	110
The Evidence for Formative Assessment	111
Approaches to Formative Assessment	120
Relationships With Other Policy Priorities	126

Chapter 5

Expertise, in Teaching and Elsewhere	135
Research on Expertise	136
Expertise in Teaching	151
Returns to Teaching Experience	158
Conclusion	161

Chapter 6

Teacher Learning	163
Content, <i>Then</i> Process	164
Changing Teacher Habits	185
Switch and Classroom Practice	200
Conclusion	205

Chapter 7

Implementation	207
An Overnight Success – in Twenty Years	208
Getting With the Program	212
A Case Study in One District	213
How Will We Know If It's Working?	219
Taking People With You	228
Managing Setbacks and Disappointments	231
The Formative Evaluation of Teaching	235
Conclusion	239

Closing Thoughts	241
-------------------------------	-----

References	243
-------------------------	-----

Introduction

The argument of this book is simple: the main job of school leaders is to improve the work performance of those they lead. The *Mabinogion* is a collection of myths and legends compiled, from earlier oral traditions, in the 12th and 13th centuries in the Welsh language. One of the legends concerns a giant named Bendigeidfran (Brân the Blessed) who is leading his army across the countryside when they come across a chasm. At the bottom of the chasm is a river, which is too deep for the soldiers to wade through, and there are rocks in the river that are so close together that the soldiers are unable to use the small boats they have brought with them. After some discussion, the giant decides that he will lay his body across the chasm, so the soldiers can walk across his back to the other side. As he lays himself down, he says, “One that would be a leader must be a bridge.”

This, it seems to me, is a powerful and fitting metaphor for leadership in schools, and neatly encapsulates the main argument of this book – that the main job of school leaders is to improve the work performance of those they lead. Other things that leaders do will, of course, have an impact on the learning of the students in the schools they lead, such as creating policies on grouping students for instruction, but the size of the impacts are small in comparison to the impact of developing the classroom practice of teachers in those schools (Educational Endowment Foundation, 2015).

There are, in addition, many things that affect educational outcomes for young people over which school leaders have little or no control. There is some evidence that the quality of school buildings can influence how much students learn (Commission for Architecture and the Built Environment, 2002), although the impact of the school environment on student achievement appears to be rather small (PricewaterhouseCoopers, 2003). On the other hand, there are many other factors, again over which leaders have little or no control, that have significant impacts on student achievement, such as poverty, teacher preparation and teacher compensation.

Of course poverty should not be an excuse for low educational achievement, but there is now a huge body of evidence that teaching children growing up in poverty is more challenging than teaching more affluent students (Dickerson & Popli, 2012; Taylor, 2010). Reducing the number of children growing up in poverty will substantially increase student achievement (Carnoy & Rothstein, 2013), but school leaders cannot do much to affect those numbers.

In many countries, and particularly in the United States, the pre-service preparation of teachers is only loosely coupled with the needs of schools, and so creating better alignment between teacher training programs and the realities of school life would also have a substantial impact on student achievement (Anderson & Stillman, 2013; Cochran-Smith & Zeichner, 2005). But again, it is hard to see how individual school leaders can do much to make university schools of education more responsive to their needs.

Closer to home, school leaders do have some influence over teacher compensation, by evaluating teachers, for example. However, often school leaders are rarely consulted about the design of teacher compensation schemes, and there is considerable evidence that these schemes are not well aligned with what matters for student achievement. For example, across the United States, compensating teachers for gaining a master's degree costs well over \$8 billion each year (Roza & Miller, 2009) even though there is little evidence that teachers with master's degrees are more effective than those without (Harris & Sass, 2007). The focus of this book is squarely on the bottom right-hand cell of the control-impact matrix shown in Table I.1.

Table I.1. Control-Impact Matrix for Schools

		Control	
		Low	High
Impact	Low	School buildings	Ability grouping
	High	Poverty	Teacher practices

In making the argument presented in this book, I have drawn together evidence from a range of sources in order to provide leaders with arguments and evidence they can use to persuade others of the importance of these ideas in their own local contexts. The result is therefore not a book designed to be read at a single sitting but rather a resource, to which I hope that readers will return from time to time for ideas to use in their own work.



[Chapter 1]

Why We Need to Raise Student Achievement, What's Been Tried and Why It Hasn't Worked

The idea that our schools need to get better has been around for a very long time, but of course before we can decide how to make our schools better, we need to be clear about what we are making them better *for*. In many countries, it would be thought absurd to design an education system without having a clear statement of the purpose of education, but in many English-speaking countries there is a widespread distrust of philosophy. This practical approach can be very useful in avoiding long debates, but there is also a danger that differences in unexamined assumptions make meaningful debate impossible. In this opening chapter, I argue that there are many reasons to educate young people, and any education system has to pull off a delicate balancing act between these. However, preparation for work deserves special attention, not because it is the most important purpose of education, but it is where things are changing fastest. In the second half of the chapter, I review some of the main measures proposed to improve education and conclude by looking at whether comparisons with other countries is helpful, using the United States as an example.

Why We Need to Raise Achievement

One of the problems with debates about education is that often people talk past each other because they differ in their beliefs about why education is important. There isn't space here for a lengthy discussion of the philosophy of education, but the main

There are a number of surprising things about this table. First, there are relatively few significant relationships of any kind – twenty-two of the thirty cells (over 70 per cent) are empty. Second, and perhaps even more surprisingly, there are as many negative as positive relationships. Third, even those significant relationships that do exist are rather difficult to understand. The fact that university courses in pedagogical content improve the effectiveness of maths teachers in primary and middle schools – but not high schools – does not seem particularly surprising, but the fact that they do not do so for reading at any level seems odd. Perhaps even more surprising is the fact that those with higher aptitude test scores are worse than those with lower aptitude scores at teaching high school maths. When one considers that with thirty cells in the table, one or two of the values would be expected to be significant just by chance, one is led to the conclusion that the table is probably just statistical noise.

In fact, this is not an isolated result. Indeed, the lack of any simple relationship between formal qualifications and the effectiveness of teachers is one of the most well-established findings in the research literature (see Table 1 in Harris, 2009, for a summary). There is evidence that what Lee Shulman (1986) called “pedagogical content knowledge” – the kind of knowledge that is needed for teaching but not for advanced work in a subject – does in fact matter, but the effects are smaller than most imagine. For example, Hill, Rowan and Ball (2005) found that primary school students taught by teachers with greater mathematical knowledge for teaching did make more progress, but the effect was small. Students taught by teachers with strong mathematical knowledge for teaching made about one month’s more progress per year than those taught by teachers with weak mathematical knowledge for teaching. This is an important difference but clearly only a small part of teacher quality. It does seem that specific pedagogical content knowledge is more important in high school, but even here, mathematical knowledge for teaching accounts for only about one-third of the variation in teacher quality (Baumert et al., 2009).

To sum up the argument so far, we know that teachers make a difference, but we don’t know what makes the difference in teachers. Physicists talk about dark matter, which is a kind of matter that they believe exists because its gravitational impact can be felt (or at least inferred from its impact on traditional matter) but is not matter in any form that we understand. We might talk about the dark matter of teacher quality in the same way.

In an article titled “Most Likely to Succeed”, Malcolm Gladwell (2008) likened picking teachers to drafting quarterbacks in the US National Football League (NFL). Performance in football at the college level predicts performance in professional football reasonably well at most positions, but at quarterback, there is little relationship between success in college and success in the NFL. At one point, it was thought that this might be because offences and defences were more complex in professional

too challenging to be solved with a single perspective (Gibbons et al., 1994) and where self-knowledge is essential to progress.

The latter is especially important because there is now extensive literature on the inadequacy of self-reports, summarised in Figure 7.1.

Only 2 per cent of high school seniors believe their leadership skills are below average (College Board, 1976) ...
... and 25 per cent of them believe they are in the top 1 per cent in their ability to get along with others (College Board, 1976).
Ninety-three per cent of Americans and 69 per cent of Swedes think they are safer drivers than average (Svenson, 1981).
Ninety-four per cent of college professors report doing above-average work (Cross, 1977).
People think they are at lower risk than their peers for heart attacks, stroke, cancer, food poisoning, etc. (Weinstein, 1980).
Strangers predict your IQ better than you do (Borkenau & Liebler, 1993).
People believe they are more accurate than their peers at accurate self-assessment (Pronin, Lin & Ross, 2002).

Figure 7.1. Why people shouldn't work on their own

The extensive research and development effort (although it would probably be more accurately described as development and research) resulted in a series of professional development resources that are being used all over the world (Northwest Evaluation Association, 2015; Wiliam & Leahy, 2014). Although at this stage, we do not have evidence from RCTs that prove the efficacy of this approach, its practical success does suggest that schools and teachers are finding the approach helpful. For example, the model is already being used in two-thirds of the primary schools in Singapore (Lee, Oh, Ang & Lee, 2014) and has been adopted state-wide in South Australia. Work in the United States is less advanced, but preliminary work in Oregon shows that when schools use these materials, there is an impact on student engagement and learning (Thum, Tarasawa, Hegedus, Yun & Bowe, 2015).

So, problem solved, surely? We had shown that formative assessment had a significant impact on student achievement, both in terms of the research literature and in practical experiments, and had found that schools, given the right tools, could reproduce the key experiences for teachers with reasonable fidelity. What we were not prepared for is how difficult it turns out to be to get schools to do what the

positive effect (Kluger & DeNisi, 1996). Leaders need to know their teachers so they know when to be critical and when to provide support. Just as important, teachers need to trust their leaders because unless the teacher believes that the leader has the teacher's best interests at heart, and unless the teacher believes the leader has credibility as a coach, the teacher is unlikely to invest the effort needed to improve practice (for more on this, see Tomsett, 2015). This means that there can be no simple recipe for effective feedback for teachers on their teaching performance, but a couple of principles derived from other research on feedback may be useful here.

The first is that feedback should cause thinking. Feedback that causes an emotional reaction, as is likely when the feedback compares an individual teacher's performance with that of other teachers, is unlikely to be helpful. Far more helpful is comparing a teacher's performance with his or her own previous performance (e.g. was this a personal best for the teacher?), which is likely to help the teacher adopt a growth mindset (Dweck, 2006). The second principle is that feedback should be more work for the recipient than the donor. The feedback event itself is likely to be relatively unimportant in improving teacher performance; what matters is the follow-up action taken by the teacher.

Activating Teachers as Learning Resources for One Another

Because, as noted above, the issue of trust between the donor and the recipient of feedback is crucial to the likely success of the feedback (Santiago & Benavides, 2009), it may be helpful to involve peers, rather than those with a formal leadership role within the school or group of schools, in providing feedback to teachers. This is particularly important where leaders have a formal role in the accountability function because it can be difficult for leaders to separate out the two roles. Specifically, it can be very difficult for leaders to ignore evidence that might be relevant to the accountability function if they are meant to be focusing on the improvement function. Even if leaders *are* able to do this, ultimately the behaviour of the teacher will depend not on whether the leader is able to separate these two roles clearly, but whether the teacher *believes* the leader is able to do so. If the teacher believes that evidence of weaknesses in practice revealed in an observation that is ostensibly intended to improve practice could also be used to affect the judgment made about the teacher's effectiveness, then the teacher is more likely to play safe so that the potential for the observation to improve practice is reduced. Where peers are involved in classroom observation, it can be particularly helpful to have a protocol for the lesson observation that makes clear that the teacher being observed:

- Specifies the focus of the observation
- Specifies the evidence to be collected
- Owns any notes made by the observer during the lesson