

The Cultural Myths and the Realities of Teaching and Learning

GRAHAM NUTHALL

Abstract:

In this article I describe evidence that most teaching follows the standard patterns of a cultural ritual that has remained largely unchanged despite many attempts to reform and change it. A series of research studies show that teachers are largely unaware of individual student learning and base their practice on the cultural ideal of the busy active classroom. Although evidence shows that differences in student learning are primarily the effect of motivation and cultural background, myths about academic ability have evolved to explain these differences and shift the blame onto students. Research is needed to replace myth and ritual with reliable evidence.

This article is about how culture shapes our understanding of both the teaching and learning process in ways that have proved extraordinarily difficult to identify and describe. If we define culture as the ways of acting, thinking and feeling that are common to the members of a society, then one of the most significant things about culture is that it becomes so much a part of ourselves that we lose awareness of how it organizes our lives. The more familiar it is, the more it is like the air we breathe, the harder it is for us to identify how it shapes what we believe and what we do.

School teaching is now an integral part of our culture. Stigler and his colleagues concluded from their detailed video study of teaching in classrooms in Japan, Germany, and the United States, that teaching, as it is commonly practised and understood, should be seen as a cultural ritual. The variations in practice that occurred in teaching within the United States were relatively minor when contrasted with the characteristic practices of Japanese teachers working within Japanese culture (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999).

Teaching, they suggested, was a ritual that we all learn through at least ten years of observation and participation as students, if not as teachers. Despite changing teacher education programmes, changing forms of classroom organisation, and many attempts to reform teaching methods, the core of the ritual remains largely unchanged, sustained by a “stable web of beliefs and assumptions that are a part of the [wider] culture” (Stigler & Hiebert, 1999, p. 87). We all become, through this common experience, experts in what it means to be a teacher and a student. As we often jokingly complain, everyone is an expert on schooling.

I have been involved in research on teaching and learning in classrooms for about 40 years, and it has taken this long for me to understand just how much of what we do in schools is a matter of cultural tradition rather than evidence-based practice. And how much of what we believe about teaching is a matter of folk-lore rather than research.

In order to explain this, I want to take you on a journey. Instead of logically explaining the case with argument and evidence, I want to trace the sequence of discoveries that I made as I observed and worked with teachers and students to try to find out how teachers shape student learning.

I want to do this, not just because it might be an interesting story, or a challenging intellectual exercise, but because I want to make it clear that so long as we are unaware of the extent to which culture determines how we practise and think about teaching, attempts at reform are likely to be ineffective. We will remain locked in a system that inevitably produces failure and inequalities.

The direction this journey took was not planned. The path was determined by my interactions with the colleagues I had the good fortune to work with, by the discoveries we made, and by the accidents and opportunities that emerged along the way.

The journey begins with a tape-recorder (1960-1968)

The journey began when I was a graduate student working on my masters thesis. I managed to persuade a group of experienced teachers to let me bring a tape-recorder into their classrooms and hang up microphones on bits of string from their light-fittings. Analysing these first tape-recordings led me to discover that, for all its apparent spontaneity, the way teachers interacted with their students followed fixed patterns and conventions. Even when a teacher was being very

sensitive to the individual needs and interests of her or his students, the interaction took place within predictable structures and rules of social interaction. Like language, teaching has its own underlying grammatical rules (Nuthall & Lawrence, 1965).

I later discovered, when I was working with a research team at the University of Illinois, that these patterns or routines were being identified by researchers in essentially the same form across different countries and languages (Simon & Boyer, 1968; Smith & Meux, 1967). Stenographic records of teaching made in the early 1900s showed that the essentials of these patterns already had a very long history (Hoetker, & Ahlbrand, 1969). Although we did not think this way at the time, we had, in fact, discovered that teaching was a kind of cultural ritual.

Learning that experience makes no difference (1968 – 1974)

The next stage in the journey occurred when I returned to New Zealand and worked with colleagues at Christchurch Teachers' College to find out how our understanding of the underlying patterns of teaching could be incorporated into the training of teachers. We tried out several ideas, including training the students to analyse recordings of their own teaching.

Working with Cliff Wright, we decided to find out if this training made a difference to the way the student-teachers taught and the effect they had on student learning. Cliff and I worked with a group of student-teachers that included some who had been trained to analyse recordings of their own teaching, a matched sample of beginning teachers who had not had the training, and a group of very experienced teachers who were experts in science teaching (in those days called nature study specialists). Each of these teachers taught the same material on the life of the black-backed seagull in their own way.

We tested the students' learning and used tape-recorders to get an exact record of how each of the teachers interacted with their students. We then related the characteristics of these interactions to what their students learned. This was a well-designed study in which we identified how different kinds of teacher questions and feedback to students affected student learning. It was reprinted as a model of this kind of correlational research (Morrison & McIntyre, 1973; Skowronek & Schmied, 1975; Wright & Nuthall, 1970).

To our surprise, however, there were no discernable differences between the experienced teachers and the beginning teachers in what

they did or what their students learned. Being an experienced expert teacher apparently made no difference.

Again, although we did not understand the significance of it at the time, we had stumbled across evidence that the basic patterns of teaching are carried out in much the same way, with much the same effects, by novices and experts alike. The underlying patterns of teaching appeared to be independent of training and experience.¹

The journey splits in two: The experimental studies (1974 – 1980)

At this point, the journey took two different directions that later joined up again. In one branch of the journey, I worked with John Church and David Hughes as they designed and carried out experimental studies of the effects of teaching on student learning. In the other branch, I set out to review the published research on teaching and write a book for teachers on what the research had to say about how to teach.

Experimental studies seemed the way to go. The problem with correlational studies like the one Cliff Wright and I did, is that you cannot be sure you are not overlooking an important aspect of teaching. The components of teaching that we identified as important could have been merely symptoms of something else much more important, something that we were not aware of.

John Church and David Hughes worked from recordings of experienced teachers, and developed scripts to structure their own teaching. These scripts ensured that they followed predetermined patterns but could still respond spontaneously within those patterns. Recordings of their scripted teaching were indistinguishable from everyday teaching.

They were both very successful sets of experiments. They discovered the effects on student learning of using different types of questions, different types of feed-back (Nuthall & Church, 1972), and different ways of managing student participation (Hughes, 1973). To my knowledge, they were the most brilliantly designed experimental studies of teaching that have ever been carried out.

But there were problems. Because of the very precise and careful way the researchers designed their experiments, they uncovered the enormous complexity of teaching. Going down that track would have meant carrying out hundreds of large scale experimental studies just to begin to study some of the different ways teachers interacted with their students. For example, John Church taught in all the available Year 5

classes in the city of Christchurch to study the interactive effects of four different types of teacher behaviour (Nuthall, 1972).

I began to realize that the end product of this kind of research would be an enormous list of experimentally validated do's and don'ts. If teachers could digest such a long list, it would turn them into robots. I realized I was following a path that satisfied the cultural rituals of the research community, but would be of little value to teachers, and probably do them harm (Nuthall, 1989).

The question started going round in my mind – what is it that teachers need to make their teaching more effective? And how could I design research methods that would meet their needs?

The other branch of the journey: A book about research on teaching (1974 – 1980)

On the other branch of my journey, I was searching the literature for studies that would throw light on the way teaching shaped student learning. I wanted to know all that research could tell me and write a book that would make this information available to teachers. What I found was a field in which the findings were extremely varied and often contradictory. I reached the conclusion that only studies that involved what teachers and students did in classrooms would ever be useful, but even then, there was surprisingly little that was both consistent and relevant (Nuthall, 1968, 1970).

At that time, I developed the belief that the essence of good research is producing results that can be replicated (Toulmin, 1961; Fodor, 1968). You cannot expect a teacher to try out a new method if the results of research on that method are just a one-off accident. And it is irresponsible to use research to change what teachers do unless you have some certainty that you know the effects on students.

Unfortunately, we almost never repeat studies to find out if the results can be reproduced in other contexts. We are always carrying out exploratory studies, or one-off studies in which we use statistical procedures to determine the value of the results. But the statistical tests that we use *assume* replicability. They do not prove it, or even establish its probability.²

At that time I knew, or had contact with, a significant number of those who were involved in classroom-based research on teaching. I managed to persuade many of them to send me their original data so that I could find out, for myself, how comparable the results of different studies actually were. It was a depressing activity. There was almost no

comparability between the different sets of data, and I felt I could put in the book little that a teacher could trust with any certainty.³

After a lot of detailed work I abandoned the book. I put the chapters I had written in a drawer and I wrote an article entitled "Is classroom interaction research worth the effort involved?" (Nuthall, 1974). It detailed the inconsistencies in the data in the different studies. And I took up painting.

The journey reaches its lowest point: Searching for the right methods

This was the lowest point in the journey. If ever there was a time in the desert, then this was it. I was not sure that research on teaching could ever produce reliable and useful results. What is more, the painstaking research that we had done by that time was being publicly attacked for being mindless dust-bowl empiricism, even fascist in intent (Delamont & Hamilton, 1976; Bolster, 1983; Garrison & Macmillan, 1984; Tom, 1985).

But these criticisms contained a very good point. Teaching is a very personal and individual thing. To be valid, research on teaching must include the subjective and personal elements of what goes on between teachers and their students (Chambers, 1992). But that just redefined the problem. How was it possible to include the subjective and the personal but still produce results that were reliable and replicable?

Meanwhile, the journals continued to publish studies that showed statistically significant results. Books were published that listed these results, glossing over inconsistencies as the errors we ought to expect in such a messy and complex area. But, there were no systematic programmes of research. Very few did more than one study, and consequently the problems of replication were never confronted (Brophy & Good, 1986).

The journey takes a new turn: Studying student learning (1978 – 1984)

A way out of this wilderness came through working with another graduate student. Adrienne Alton-Lee was an experienced teacher who was concerned about how little she knew about student learning. She knew how to manage a class. She knew how to interest and engage the students in learning activities. What she did not know was why a particular student did not learn a particular concept when other students did. She designed her PhD thesis so that she could follow the learning experiences of individual students. With two other observers,

she made continuous written records of the behaviours of three students in a Year 4 class, during a unit on environmental issues (Alton-Lee, 1984).

One of the major contributions that Adrienne made was her invention of what she called the "item-file". This was a data file made up of the records, both objective and subjective, of every experience a student had that related in any way to the learning of a single concept. Each item-file was, as far as possible, the complete life history of a single concept in the mind and experience of an individual student. Please note that here (and later in this article) I am using the term "concept" as a short-hand for all the different kinds of knowledge, understandings and skills that are significant in areas like science, social studies, and mathematics.

Because these item files contained every detail of the students' experiences, they gave a sense of being as valid and sensitive to the real lives of students as any research in this field had ever been. And analysing them produced some very significant insights into student experience and student learning.

Although Adrienne's thesis was, in many ways, a brilliant success, and some of the results were published in a well-respected international journal (Nuthall & Alton-Lee, 1990), I had doubts about whether the results from just three students in one classroom were enough. So the journey, for the next eight to ten years was spent carrying out three increasingly sophisticated replications of Adrienne's original design. Roger Corbett helped enormously by designing individual miniature microphones for each student to wear. He worked out ways of mounting multiple sets of small video-cameras to supplement the live observations. We were now able to study every kind of classroom activity in continuous detail.

The journey continues through three replications (1984 – 1990)

The first thing that became apparent from this very detailed data on the experiences of individual students was how little teachers knew about what was going on in their classrooms. We found that even live observers keeping continuous written records of the behaviours of individual students missed up to 40 percent of what was recorded on the students' individual microphones (Alton-Lee, Nuthall, & Patrick, 1993).

We began to realize that students lived in a personal and social world of their own in the classroom. They whispered to each other and

passed notes. They spread rumours about girlfriends and boyfriends, they organized their after-school social life, continued arguments that started in the playground. They cared more about how their peers evaluated their behaviour than they cared about the teacher's judgement. Within this pervasive peer culture, sexism and racism were alive and flourishing even when the teacher actively promoted fully inclusive learning activities (Alton-Lee, Nuthall, & Patrick, 1987; Alton-Lee, Nuthall, & Densem, 1990).

I also came to realise that student learning is a very individual thing. Students already know at least 40-50 percent of what teachers intend them to learn. Consequently they spend a lot of time in activities that relate to what they already know and can do. But this prior knowledge is specific to individual students, and the teacher cannot assume that more than a tiny fraction is common to the class as a whole. As a consequence, at least a third of what a student learns is unique to that student, and the rest is learned by no more than three or four others (Nuthall, 1999a).

Understanding the function of the standard rituals of teaching

I now began to understand the function of the standard patterns or routines of teaching, and why they had such control over teachers' behaviour. In order to manage a class of 25 to 35 students, all of whom have different knowledge, skills, interests and motivations, teachers have to focus on the performance of the class as a whole. It is impossible to focus on the individual learning of any one student for more than very brief periods.

Within these standard patterns of whole-class management, students learn how to manage and carry out their own private and social agendas. They learn how and when the teacher will notice them and how to give the appearance of active involvement. They get upset and anxious if they notice that the teacher is keeping more than a passing eye on them.

One way to understand this is to think of the class as an orchestra following a musical score. So long as everyone knows their parts, the whole works together effectively. The coherent sound of the whole orchestra makes it extraordinarily difficult to separate out the sounds of individual instruments. If teaching is like conducting an orchestra, then it must be primarily about group management and must follow predictable patterns, so that both teacher and students know how to interact with each other.

Discovering how students learn from classroom experience (1990-1995)

The next stage in my journey occurred during 1990 when I took study leave to work full-time on analysing the data from the three studies. My primary focus was on the factors that affected student learning. For nine months of that study leave, I did nothing else but analyse that data. I used everything I knew about formal and informal data analysis, producing endless computer printouts and pads full of notes. But gradually it dawned on me that for all the richness of the data we had obtained, the old problem still remained. Adrienne's PhD results did not reappear, and each of the new studies produced what looked like different results. The relationships between teacher and student behaviours, activity structures, resources and student achievement seemed unique to each study.

I kept thinking that if I added in more variables, included more detail, some kinds of patterns would emerge. But they didn't. At this point in the journey, it seemed that unless we could get inside the minds of students, we would never understand exactly how learning occurred.

But I decided, just to see what it would look like, to take one student's learning of one concept, and see if I could make sense of what was going on in that student's mind. It was a boy, code-named John, and it was an item-file about recent migration to New York.

As I followed the detail of his experiences, the things he saw, heard, read, said, wrote and joked about, the obvious started to dawn on me. Learning is usually a progressive change in what we know or can do. What creates or shapes learning is a sequence of events or experiences, each one building on the effects of the previous one. An event at one point in the sequence will have a different effect from the effect the same event would have had if it had occurred at another point in the sequence. None of the methods of data analysis that I or anyone had used in this kind of research allowed for this possibility.

It also dawned on me that what is important about a student's experiences is the information that she or he can extract from those experiences. It is less important what that student is doing, or what resources the student is using, or what are any of the other contextual aspects of the experience. What matters is the sense the student is making of the experience.

In hindsight, this seems very obvious, but it shows how caught up you can get in the culture of the research community. In the desire to focus on general variables of theoretical significance that are relevant

across many different contexts, we have been blind to the significance of the particular.

So I developed a way of identifying the information that students extract from their experiences, and a way of classifying the content, timing and sequencing of that information. To give a simple example, if a student is expected to learn that Antarctica is the driest of all the continents, there are many different kinds of relevant information. The student might read or hear about the kinds of storms that occur in Antarctica. The student might see a video showing how the members of an expedition in Antarctica coped with the weather conditions. Or the teacher might tell the students that, despite what everyone thinks, there is less rainfall in Antarctica than there is in the Sahara desert. Each of these kinds of information is relevant, but in different ways, and to a different extent.

My study of John showed that when he learned a concept, he had experienced at least three times the complete set of the information he needed to fully understand the concept. Each complete set of information could come from a single experience, or could be made up of different kinds of information coming from different experiences at different times. Without the three sets of complete information, he did not learn the concept. It was as simple as that.

Predicting what students will and will not learn

This provided the basis for a set of rules for determining when a student had encountered sufficient information to understand, learn and remember a concept. I applied these rules to the experiences of the other three students in John's class, made some minor adaptations and then used them to predict exactly what concepts the students in the other two studies would learn or not learn.

These rules worked. They successfully predicted, with about 80 to 85 percent success, the learning of nearly 500 concepts through the experiences of eleven children, in the three different classrooms. It involved a lot of detailed work but it was replication of the kind I had never seen before. It was at last a logical method that fitted the problem (Nuthall, 1999a; Nuthall & Alton-Lee, 1993, 1994).

This was the first major discovery that I had made. There was nothing in anybody else's research that looked anything like this. Maybe now the journey was headed in the right direction. I had found a way of relating student experience to student learning. But I still had to connect student experience to teaching.

Discovering how students answer tests

Another significant discovery occurred at this time. We had been very careful to try to find out exactly what each student knew and had learned. We not only used very carefully developed and administered paper and pencil tests, but also extensive individual interviews with students that explored their learning experiences and their knowledge and understanding in greater depth. This allowed us to look closely at how the students answered paper and pencil tests, and how they responded in interviews.

Out of this analysis came the understanding that testing is like interviewing. Both depend on the relationship between tester and student. How the student responds depends on the nature of that relationship and the extent to which the student and the tester share the same perceptions of what the test or interview is about. Despite its apparent objectivity, there is nothing more or less objective about a test than there is about an interview. There is just a different kind of relationship between the tester and the student (Nuthall & Alton-Lee, 1995).

I also came to understand that what a student knows and can do is a coherent body of beliefs and understandings that is not the same as an adult's view of the world, but is nevertheless logical and consistent. Distinctions between facts, concepts, principles, generalizations and procedures exist more in the theories of researchers than they do in the minds of the students. Knowledge is more like a continuous landscape rather than a set of discrete countable objects. It cannot be sensibly represented by numbers. This led to the conclusion that the scores that students get on standard paper and pencil tests are primarily the result of the students' motivations and cultural background, and only secondarily about what the student knows or can do.

To get a sense of this, imagine a hot nor-west afternoon in a classroom with all the windows shut to keep the wind from blowing everything around. A teacher the class doesn't particularly like, or some stranger, has just handed out tests and tried to persuade the students to do their best. Imagine a student starting off with the first few easy questions, then as the room gets hotter and her clothes get stickier and the questions get harder, wondering what the hell she is doing this for. She starts doodling instead of filling in the answers and starts to realize that no one she knows gives a damn about what she does on the test. She sucks her water bottle, draws a few more faces to keep boredom at bay, and waits for the test to be over and real life to begin again.

If this is a fair representation of the experience of a significant minority of students (and my own observations suggest it is), then research on teacher and school effectiveness and research on standards of student achievement (especially the large international studies such as TIMSS and PISA), because they depend on paper and pencil tests that have little personal significance for many students, will frequently produce invalid results.⁴

The journey takes a dramatic turn: Identifying the role of ability (1995 – 1998)

But the journey gets more exciting and disturbing at this point. Adrienne and I carried out two further studies that focused on gender and curriculum issues. These studies allowed me to develop and extend the prediction rules to cover further kinds of learning and to test their predictive validity with older students learning different kinds of concepts. Again they worked just as well. The procedure was now objective enough to be computerized. The tally was now 1100 different concepts, learned or not learned by 21 students in 6 different classrooms (Nuthall 1999a; 1999b).

However, when somebody asked me to explain what effect ability had on learning, I had to reply that as far as I could tell, from our data, none at all. The same rules worked equally well for both the more able and the less able students. But our data did show that the more able students started with more background knowledge and ended up learning more than the less able students.

This started me comparing the experiences of different students when they were engaged in the same activities. What I found was that a large proportion of each student's significant learning experiences were either self-selected or self-generated, even in quite traditional classrooms.

The more able students talked more amongst themselves about relevant content. They asked more questions and persisted with problems for a longer time. They seemed to be more interested, more persistent, and less likely to be distracted. There was no evidence that they found the tasks easier, or had fewer difficulties. There was no evidence that their minds processed their experience differently. The difference was in the way they managed their involvement in classroom activities, and in the advantage they gained from having more relevant background knowledge (Nuthall, 1999, April; 2001; 2001, April).

It was Piaget's work that provided the metaphor that helped me understand the problem. Piaget was originally a biologist, and began his work on the development of mind in the belief that as the digestive system processes and extracts what it needs from food, so the mind extracts, digests, and grows from what it needs from experience.

Within reasonable limits, the learning process, like the metabolic processes that take place in the body, is universal across our species. How the metabolic process actually works in our bodies depends on what we eat and drink. So it is with the learning process. Although it always operates in the same way, we can, as it were, choose our mental diet, feeding or starving our minds of its essential nutrients.

So those students whose backgrounds provide them with the cultural knowledge and skills to use the classroom and its activities for their own purposes, learn more than those who dutifully do what they are told but do not want, or know how, to create their own opportunities. Differences in ability are more likely to be the product of differences in classroom experiences than the other way round.⁵

The journey again splits (1990 – 2000)

At this point, my journey again split into two parallel directions. Along one branch of the journey, I started to look more closely at the rituals that make up classroom life. If Vygotsky is right, then students' minds develop by internalizing the social rituals of classroom life. If we want to understand how classroom experiences shape students' thinking and the way their minds process experience, we need to know exactly what these social rituals or routines are. I already had a clear understanding of the recurring patterns of whole class activities, but needed to discover the recurring patterns of small group and individual activities (Nuthall, 1999b; 2000a).

Learning from experienced teachers

The other branch of the journey came from my experiences with graduate students. I developed a course on research on teaching for experienced teachers and lecturers. It was a practical course and I learned a lot from these students about the nature of teaching and how experienced teachers think about their teaching.

Two assignments in particular provided significant insights. In one assignment, I asked the graduate students to interview fellow teachers about how they knew when their teaching was going well. The teacher

was asked to describe an actual incident when their teaching went especially well, then describe how they knew it was going well (based on a procedure developed by Jackson, 1968).

What was fascinating about that assignment was that it always (over more than a dozen years) produced the same results, with minor variations depending on the age of the students in the teachers' classes. Almost every teacher knew their teaching was going well from signs of students' engagement. It was the look in the students' eyes, the questions they asked, the fact that they didn't stop talking about the topic or problem when they left the classroom. In short, by the feel and sounds of interest and focused busyness. In most teachers' minds, the criteria for successful learning were the same as the criteria for successful management.

The second assignment I set my class was much more difficult. I asked them to carry out research on their own teaching. They could take any small teaching activity or procedure, measure what three or four students learned from that activity, and relate, in some direct way, what happened in the activity to what the students experienced and learned.

Despite instructions and discussion with me, many of the graduate teachers and lecturers carried out the assignment in the same way. They recorded what they and their students did together and created an outcome test to measure what the students learned. They then related the students' scores on the test to indicators of how engaged the students were in the teaching or learning activity. And they found that the more engaged the student, the higher the score on the test.

It was difficult for me to explain and hard for these experienced teachers to understand that I wanted them to do more than that. The problem with that kind of study is that it is not directly about learning. It is about classroom management. In order to relate teaching to learning, you need to know how a student's knowledge or skills have changed and what specifically caused that change to occur. Being busy is not a cause of learning unless you know exactly what information or knowledge the student is getting out of being busy. To give a simple example, it is not enough to say that a student was busy reading unless you also describe what the student was reading and how that related to what the student learned or failed to learn.

What I learned from these graduate students and the difficulties we had communicating with each other was that the practice of teaching, as we commonly understand and talk about it, is not about learning.

The focus of teachers' thinking when they are planning and carrying out their role in the classroom is keeping students busily engaged in activities that produce some tangible product that may or may not reflect, in some indirect and unobservable way, student learning (see, for example, Clark & Peterson, 1986; Leinhardt & Greeno, 1986; McNair, 1978).

As a deputy principal of a large high school said in his evaluation of the course: "I realized I had not really consciously thought about what effective teaching is ... I had made the assumption that because I was teaching, the children were learning."

Identifying the rituals of individual and group activities (1998 - ?)

At this point, the two branches of my journey joined up again. I had become involved with a group that was looking at the cultural and historical roots of classroom practices (Nuthall, 1997). In my attempts to identify the routines and rituals of the classroom, I looked closely at how teachers managed the activities they designed.

What was immediately apparent was that teachers do not talk to students about learning or thinking. They talk about paying attention and not annoying others. They talk about the resources the students will need to use, about how long the activity should take and what will happen if it is not finished on time.

When you listen to students they talk about the same things. They are constantly comparing how much have they done. How long will it take, do the headings have to be underlined, where did you find that answer, do you have to write it all out, does it have to be finished for homework? (Nuthall, 2001; 2001, August).

Homework is a classic example of the way learning activities are thought about. Everyone (parents, politicians, teachers, some students) believe it is important. There is debate about how much and how often it must be done, but giving no homework raises suspicions about whether a teacher is doing a good job. However, nobody talks about what or how a student learns from doing homework. Nobody cares if the parents do most of it, not even the parents themselves. The point is that homework, like most classroom activities, results in a more or less beautifully presented product. Of course there is supporting talk about it being good to keep students busy and away from the television set, but not about the need for the specific learning that a specific homework project might produce.

Where has the journey taken me?

Let me pause now to review where the journey has taken me to this point. The appropriate image might be standing on an isolated hill, distant from the standard ways of doing research, distant from the standard ways of understanding teaching and learning, doubting the value of the accumulating bodies of published research on teaching effectiveness, academic ability, and achievement. But all the while, motivated by a series of discoveries that seemed to be getting me closer to a real understanding of how classroom teaching relates to learning.

First, you will have noticed a major change of direction in the middle of the journey. To be precise, at the time of Adrienne Alton-Lee's PhD thesis, my work with her involved a change from focusing on teaching to looking more closely at student learning.

My primary interest was still in relating teaching to learning, but we had found that students live in a different world from teachers and testers and researchers. If we are to understand how teaching relates to learning, then we have to begin at the closest point to that learning, and that is student experience.

Second, I had come to understand that there are very good reasons why teaching and teacher-managed classroom activities follow very predictable patterns that are only indirectly related to student learning. This is because teachers are very largely cut off from information about what individual students are learning. Because of the numbers of students that teachers have to manage simultaneously, and because of the individuality of student learning, teachers must rely on routines and rituals that are commonly believed to be good for students. For example, many teachers make use of the class brainstorm as a routine for finding out what students know. In a brainstorm, students are invited to contribute any or all of the ideas they have on the topic. Typically, a few students contribute the majority of the ideas, a few more students contribute one or two ideas, and most students are silent. The mesmerizing power of whole-class routines, like the brainstorm, is such that most teachers, and most observers, come away feeling they know what all the students know. Even though most of the students said little or nothing.

To make matters worse, when a teacher uses a brainstorm to review and evaluate what the students have learned from an activity, and the teacher is unaware that most of what the contributing students learned, they learned from their own self-created activities, the teacher will come

away with an entirely false impression of the role she or he played in the students' learning.

However, so long as these routines and rituals are recognized by the profession and the society at large as the right way to run classrooms, and students have learned to expect and play the reciprocal roles that these rituals require, then teachers do not have to pay more than passing attention to what is going on in the students' minds. Which is why attempts to reform teaching and to get teachers to teach in significantly different ways have a very long history of failure (cf. Williams, 1966). As Cuban (1998) has demonstrated, teachers are more likely to reform the reforms than be reformed by them.

Identifying the web of supporting myths: The role of ability

I also came to understand that there is a web of supporting beliefs or myths that justify the way these rituals are played out. Possibly the most significant of these is the concept of academic ability. When the teacher manages the classroom so that all the students are, as far as the teacher can see, busily engaged in appropriate activities, what can explain consistent differences in what the students produce. The teacher is doing everything right, so why do some students fail to respond?

The concept of inherent ability or intelligence appropriately transfers responsibility away from the teacher's management of classroom activities. Learning is said to be a function of ability. Our data, however, show that differences in what students learn, and differences in what they do on tests, are both strongly affected by differences in how they engage with classroom and testing activities. In both cases, these are a function of their motivation and the extent to which they share the purposes and culture of the teacher or tester (Nuthall, 2001, April).

This is why ability tests are good predictors of results on school achievement tests, and relate to the kinds of work that students produce in class. It all fits neatly together, and provides the basis for depriving some students of resources, and lowering the quality and content of what they are taught. The fact that differences in motivation and culture are also related to differences in ethnic background and family resources, means that the myth explains failure while protecting the real causes of that failure. It also provides the context in which students acquire false but self-fulfilling beliefs about their own ability (Nuthall, 1999b; Oakes, Wells, Jones, & Datnow, 1997).

What sustains the rituals and their supporting myths?

If the standard models of teaching and the network of myths that support them have such a long history and have such a powerful hold over the way we organize, run and think about schooling, what is it that sustains and promotes them?

First, our teacher education systems still reflect an apprenticeship model of training in which the practices and beliefs of experienced teachers are taken as the ideal to be imitated. Beginning teachers' preoccupation with classroom management leads them to focus on the surface features of the classrooms of experienced teachers. They strive to imitate the ease and fluency with which the experienced teacher manages her or his class.

The reflective teacher model that has come to dominate thinking about teacher education and professional development for teachers, also serves to sustain the standard routines and rituals of teaching. It presupposes that the reflective teacher has valid information about what is happening in her or his classroom and what the students are learning.

Most of the ways we evaluate teachers are based on the same standard models. ERO officers are experienced teachers who base their judgments on the expected routines and rituals of the busy active classroom. These are the routines and rituals that parents understand and expect to see. Those teachers who perform these routines with the most enthusiastic participation of their students are the ones who are given awards for the quality of their teaching. Others are invited to watch and copy what these award winning teachers do.

More significantly, the standard routines and rituals of teaching are strongly supported by a considerable body of academic research. Ethnographic studies that are based on teachers' perceptions and self-reports of their own teaching serve to elaborate and justify the standard routines and the myths that support them. Empirical studies of the practices of the best teachers perform the same function. In these studies, the best teachers are selected by asking school principals or other teachers to identify those who have the reputation of being the best exponents of the standard routines and rituals. The best descriptions of exactly what these routines consist of, and what teachers believe about these routines, are to be found in the reports of this kind of research (see, for example, Ramsay & Oliver, 1995). The point is that, in none of these examples, is there any direct reference to the learning of students.

It is also significant that I had to *discover* how students learn in classrooms, despite a large number of textbooks on student learning. These textbooks contain theories of learning and descriptions of how to apply them to teaching. With the exception of applied behaviour analysis (which is primarily about the management of classroom behaviour), none of the theories in the textbooks is based on research on students' experiences in classrooms. It proved impossible to relate them in any sensible way to our data. At best what these theories do is elaborate the talk about what teachers and researchers imagine goes on in classrooms. It is also interesting how little they say about memory, yet I have found that understanding memory is critical to understanding classroom learning (Nuthall, 2000a; 2000b).

The system in which learning does not matter

We seem to have created a system in which it does not really matter if students learn or fail to learn from their classroom experiences. No one is held accountable if students fail to develop a scientific understanding of the physical world after at least six years of systematic science teaching. It does not really matter if students leave school knowing almost no history or geography or economics or political science, or if students enter tertiary study with only basic arithmetic skills.

What matters in the system is that numbers representing test results are recorded in the school's information management system. Although nobody knows how these numbers or marks relate to what students actually know or can do, they are used as the primary evidence that we are concerned about student learning. More than that, many seem to believe that the more of these numbers we put in the records, the better the teaching will become.

Within such a system it is safe to talk about new ways of teaching without ever having to establish how they might affect student learning. The theory of learning styles is a classic example. It is strongly supported by many teacher educators without any evidence that it relates to student learning. As far as I can tell, those teachers who do embrace it, still engage in the established routines of teaching, but talk as though they have a new sensitivity to the learning of individual students.

Let me make it clear that this article is not intended to blame the many sensitive and creative teachers who work strenuously to promote the well-being of their students. There are very good reasons why most teachers must stay with the standard routines and rituals of teaching.

Not only is the system and its supporting myths designed to make it almost impossible for them to do or think otherwise, but it is not clear that one teacher can possibly be responsible for the learning of 25 to 35 students simultaneously. No one has ever given clear empirical evidence that a teacher can.

My major concern is with the failure of the educational research community to understand the problem that we face. As you will be able to tell, I believe in the need to include both the subjective elements of human experience and the objective analysis of patterns of behaviour in the same studies. The major problem we have to deal with is disentangling the patterns that are culturally determined and hence within our control, and the patterns that are biologically controlled, and mostly beyond our control.

At the present time we need research that focuses on the realities of student experience and the learning that results from that experience. That does not just mean "giving students a voice". It means developing a precise and accurate, if you like, scientific or replicable, account of the realities of their experiences. In my view, the truth lies in the detail. Every generalization we make, every conclusion we draw, must be true of every individual. That is what ethically responsible research in education must be like. And, as my journey indicates, it may take a very long time and a lot of data to begin to begin to discover the reality beneath the cultural rituals and myths by which we live, and evaluate them for what they are (Nuthall, 1989).

Let's return to that classroom on that nor'-west afternoon. Someone has taken his only pen away from the class clown. It's being passed around behind his back so that every time he thinks he's worked out where it is, it is somewhere else. Most of the class who know what's going on can barely restrain themselves from bursting into uncontrolled giggles. The teacher thinks they are smiling in appreciation at something she has just said. In a minute or two, she will gather up the test papers and send them off to get scored, one or two covered in doodles, others, like the class clown's, largely blank. These scores will be entered into machines where they will be transformed in complex and sophisticated graphs and tables that politicians and newspaper editors will use to berate and praise – you know that story.

So long as we do not disentangle the myths from the reality we will remain slaves to a system that inevitably produces failure and inequalities. The key lies in promoting research that allows us to stand back from the system and discover exactly how it functions.

Notes

1. More recent research has found inconsistent and conflicting relationships between teacher experience and student achievement, cf. Dewey, Husted, & Kenny (2000), Verstegen & King (1998).
2. There is not space in this article to justify this claim, but briefly, most of the statistical tests that are commonly used in research on teaching are based on the assumption that the variance of the scores obtained on the measures used in a study provide a good estimate of the variance in the population at large. In other words the data obtained in the study are used as a proxy for the kind of data that would be obtained if we studied the entire population. Since we never replicate studies, we never test this assumption. Further, it is never clear in studies of the effects of teaching, what kind of population we are sampling, or what it might mean to obtain a random sample of, for example, teaching methods or behaviours.
3. In particular, I am grateful to Ned Flanders who generously gave me copies of all the data he obtained in a very extensive series of studies of classroom interaction and student achievement.
4. There is surprisingly little research on the validity of standardized test items in which students have been asked to explain why they gave, or selected, specific answers. But see Harlow (2000), Tamir (1990), Wang (1998) for suggestive studies, and Eley & Caygill (in press) for analysis of the data from the NEMP results. The NEMP studies are an important exception because of the variety of formats they use, and the use of one-to-one testing.
5. There is interesting supporting evidence for this view from studies in which teachers have been asked to rate their students on ability, motivation, and achievement. Typically their ratings do not distinguish motivation from ability or achievement (see for example, Goldenberg, Gallimore, Reese, & Garnier, 2001).

References

Alton-Lee, A. G. (1984). *Understanding learning and teaching: An investigation of pupil experience of content in relation to immediate and long-term learning*. Unpublished doctoral dissertation, Education Department, University of Canterbury.

Alton-Lee, A. G., Nuthall, G. A., & Densem, P. (1990). I only think of men ... I don't think of women. *set: Research Information for Teachers*, No. 2, Item 16.

Alton-Lee, A. G., Nuthall, G. A., & Patrick, J. (1987). Take your brown hand off my book: Racism in the classroom. *set: Research Information for Teachers*, No.1, Item 8.

Alton-Lee, A. G., Nuthall, G. A., & Patrick, J. (1993). Reframing classroom research: A lesson from the private world of children. *Harvard Educational Review*, 63(1), 50-84.

Bolster, A. S. (1983). Toward a more effective model of research on teaching. *Harvard Educational Review*, 53, 294-308.

Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching*. (3rd ed.) (pp. 328-375). London: Collier Macmillan.

Chambers, J. H. (1990). Unacceptable notions of science held by process-product researchers. In R. Page, et al. (Eds.), *Philosophy of Education, 1989: Proceedings of the 44th annual meeting of the Philosophy of Education Society* (pp. 81-95). Normal, Illinois: Philosophy of Education Society, Illinois State University.

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.), (pp. 255-296). New York: Macmillan.

Cuban, L. (1998). How schools change reforms: Redefining reform success and failure. *Teachers College Record*, 99(3), 453-477.

Delamont, S., & Hamilton, D. (1976). Classroom research: A cautionary tale. In M. Stubbs & S. Delamont (Eds.), *Explorations in classroom observation* (pp. 3-20). New York: Wiley.

Dewey, J., Husted, T. A., & Kenny, L. W. (2000). The ineffectiveness of school inputs: A product of misspecification? *Economics of Education Review*, 19(1) 27-45.

Eley, L., & Caygill, R. (in press). One test suits all? An examination of differing assessment formats. *New Zealand Journal of Educational Studies*.

Fodor, J. A. (1968). *Psychological explanation: An introduction to the philosophy of psychology*. New York: Random House.

Garrison, J. W., & Macmillan, C. J. B. (1984). A philosophical critique of process-product research on teaching. *Educational Theory*, 34, 255-274.

Goldenberg, C., Gallimore, R., Reese, L., & Garnier, H. (2001). Cause or effect? A longitudinal study of immigrant Latino parents'

- aspirations and expectations and their children's school performance. *American Educational Research Journal*, 38(3), 547-582.
- Harlow, A. (2000). *Student responses to international assessment items in science: a TIMMS example*. Unpublished MEd thesis, University of Waikato.
- Hoetker, J. & Ahlbrand, W. P. (1969). The persistence of recitation. *American Educational Research Journal*, 6(2), 145-167.
- Hughes, D. C. (1973). An experimental investigation of the effects of pupil responding and teacher reacting on pupil achievement. *American Educational Research Journal*, 10(1), 21-37.
- Jackson, P. W. (1968). *Life in classrooms*. New York: Holt, Rinehart, Winston.
- Leinhardt, G. & Greeno, G. (1986). The cognitive skill of teaching. *Journal of Educational Research*, 78(2), 75-95.
- McNair, K. (1978). Capturing in-flight decisions: Thoughts while teaching. *Educational Research Quarterly*, 3(4), 26-42.
- Morrison, A., & McIntyre, D. (Eds.) (1973). *Social psychology of teaching*. Harmondsworth, England: Penguin Books.
- Nuthall, G. A. (1968). Studies of teaching, II: Types of research on teaching. *New Zealand Journal of Educational Studies*, 3(2), 125-147.
- Nuthall, G. A. (1970). A review of some selected recent studies of classroom interaction and teaching behavior. In J. J. Gallagher, G. A. Nuthall, & B. Rosenshine, *Classroom observation* (pp. 6-29). American Educational Research Association Monograph Series on Evaluation. Chicago: Rand McNally.
- Nuthall, G. A. (1972). *The University of Canterbury Teaching Research Project. Classroom Interaction Newsletter*, 7(2), 3-13.
- Nuthall, G. A. (1974). Is classroom interaction research worth the effort involved? *New Zealand Journal of Educational Studies*, 9(1), 1-17.
- Nuthall, G. A. (1989). *How can research on teaching help teachers*. NZARE 1989 Conference Proceedings: Keynote Addresses, pp. 33-60, Wellington.
- Nuthall, G. A. (1997). Understanding student thinking and learning in classrooms. In B. J. Biddle, T. L. Good, & I. F. Goodson (Eds.), *The International Handbook of Teachers and Teaching* (pp. 681-768). Dordrecht: Kluwer Academic Publishers.

- Nuthall, G. A. (1999a). The way students learn: Acquiring knowledge from an integrated science and social studies unit. *Elementary School Journal*, 99(4), 303-341.
- Nuthall, G. A. (1999b). Learning how to learn: The evolution of students' minds through the social processes and culture of the classroom. *International Journal of Educational Research*, 31(3) (whole issue).
- Nuthall, G. A. (1999, April). *How students learn: The validation of a model of knowledge acquisition using stimulated recall of the learning process*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal.
- Nuthall, G. A. (2000a). The anatomy of memory in the classroom: Understanding how students acquire memory processes from classroom activities in science and social studies units. *American Educational Research Journal*, 36(2) 247-304.
- Nuthall, G. A. (2000b). *How children remember what they learn in school*. Wellington: New Zealand Council for Educational Research.
- Nuthall, G. A. (2001). Understanding how classroom experience shapes students' minds. *Unterrichts wissenschaft: Zeitschrift für lernforschung*, 29(3), 224-267.
- Nuthall, G. A. (2001, April). *Student experience and the learning process: Developing an evidence based theory of classroom learning*. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle.
- Nuthall, G. A. (2001, August). *Bridging the gaps: The interactive effects of instruction and social processes on student experience and learning outcomes in science and social studies activities*. Address to the European Association for Research on Learning and Instruction 9th Biennial Conference, Fribourg, Switzerland.
- Nuthall, G. A., & Alton-Lee, A. G. (1990). Research on teaching and learning: Thirty years of change. *Elementary School Journal*, 90, 547-570.
- Nuthall, G. A., & Alton-Lee, A. G. (1993). Predicting learning from student experience of teaching: A theory of student knowledge acquisition in classrooms. *American Educational Research Journal*, 30(4), 799-840.
- Nuthall, G. A., & Alton-Lee, A. G. (1994). Understanding how pupils learn. *set: Research Information for Teachers, No.2, Item 3*.

- Nuthall, G. A., & Alton-Lee, A. G. (1995). Assessing classroom learning: How students use their knowledge and experience to answer classroom achievement test questions in science and social studies. *American Educational Research Journal*, 32(1), 185-223.
- Nuthall, G. A., & Church, R. J. (1972). Experimental studies of teacher behaviour. In G. Chanan (Ed.), *Towards a Science of Teaching* (pp. 9-25). London: National Foundation for Educational Research.
- Nuthall, G. A., & Lawrence, P. J. (1965). *Thinking in the classroom*. Wellington: New Zealand Council for Educational Research.
- Oakes, J., Wells, A., Jones, M., & Datnow, A. (1997). Detracking: The social construction of ability, cultural politics, and resistance to reform. *Teachers College Record*, 98(3), 482-510.
- Ramsay, P., & Oliver, D. (1995). Capacities and behaviour of quality classroom teachers. *School Effectiveness and School Improvement*, 6, 332-366.
- Rosenshine, B. (1970). *Teaching behaviours and student achievement*. Windsor, UK: National Foundation for Educational Research.
- Simon, A., & Boyer, G. E. (Eds.) (1968). *Mirrors for behavior: an anthology of classroom observations instruments*. Philadelphia: Research for Better Schools Inc.
- Smith, B. O., & Meux, M. O. (1967). *The strategies of teaching*. Urbana, Illinois: University of Illinois Bureau of Educational Research.
- Skowronek & Schmied (Eds). (1975). *Forschungstypen und forschungs-strategien in der erziehungswissenschaft*. Hamburg: Hoffman & Camper.
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMMS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan and the United States*. Washington: National Center for Education Statistics, US Department of Education, NCES 99-074.
- Stigler J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: The Free Press.
- Tamir, P. (1990). Justifying the selection of answers in multiple choice items. *International Journal of Science Education*, 12(5), 563-573.
- Tom, A. R. (1985). Rethinking the relationship between research and practice in teaching. *Teaching and Teacher Education*, 1, 139-154.

- Toulmin, S. E. (1961). *Foresight and understanding: An inquiry into the aims of science*. Bloomington: Indiana University Press.
- Verstegen, D. A., & King, R. A. (1998). The relationship between school spending and student achievement: A review and analysis of 35 years of production function research. *Journal of Education Finance*, 24(2), 243-262.
- Williams, J. D. (1966). Method-reversion: The problem of sustaining changes in teacher behaviour. *Educational Research*, 8 (2), 128-133.
- Wright, C. J., & Nuthall, G. A. (1970). Relationships between teacher behaviors and pupil achievement in three experimental elementary science lessons. *American Educational Research Journal*, 7(4), 477-491.

The author

Graham Nuthall is Professor Emeritus in Education at the University of Canterbury. His research on teaching and learning in classrooms has been published in books and articles in the United States, Europe, Britain and Australia, and translated into several languages. He recently received the McKenzie award and the Royal Society Science and Technology medal for his research.