

The Cultural Myths and Realities of Classroom Teaching and Learning: A Personal Journey

GRAHAM NUTHALL

University of Canterbury

In this article, I argue that classroom teaching is structured by ritualized routines supported by widely held myths about learning and ability that are acquired through our common experiences as students. These ritualized routines and supporting myths are sustained not only by everyone's common experience of schooling, but by teacher education practices, the ways we evaluate teachers' classroom performance, and many common types of educational research. My own research on teaching over the last 45 years has produced a number of apparently contradictory and puzzling findings that have progressively led me to understand the nature and power of these routines and myths. While ritualized routines are necessary to allow a teacher to manage the experiences of 20–30 students simultaneously, they also explain why individual student experience and learning remain largely invisible to teachers. The problem is to find ways to stand outside the ritualized routines and myths to identify how they control what we perceive, believe, and do about reforming teaching and learning.

As part of the Third International Mathematics and Science Study (TIMSS), Stigler and his colleagues carried out a detailed video study of teaching in a sample of classrooms in Japan, Germany, and the United States (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). One of their major conclusions was that teaching, as it is commonly practiced and understood, should be seen as a cultural ritual. The variations in the practice of teaching that occurred within the United States were relatively minor when contrasted with the characteristic practices of Japanese teachers working within Japanese culture. Teaching, they suggested, was a ritual that we all assimilate through at least 10 years of participation as students. Despite changing teacher education programs 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.

If we define culture as the customary ways of acting, thinking, and feeling that are common to the members of a society and sustain their relationships (Bohannon, 1995), 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 shapes our perceptions and organizes our lives.¹ “Much of what we do every day is an indirect expression of submerged cultural understandings (or hidden assumptions) that permeate every dimension of our beliefs and attitudes” (Spindler, 1999, p. 468).

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. What I want to recount in this article is the story of my own research journey and how, through nearly 45 years of research on teaching and learning in school classrooms, I have slowly become aware of how much of what we do in schools and what we believe about teaching and learning is a matter of cultural routines and myths. What is more, much of the research on teaching and learning in classrooms is itself caught up in the same rituals and myths and sustains rather than challenges these prevailing beliefs.

This article takes the form of a narrative. I have used this form because it allows for the interweaving of the many different strands of evidence accumulating through time that have led to my present awareness. What holds the story together is the progressive interaction of evidence, sometimes deliberately, sometimes accidentally encountered through research on teaching, through my own teaching, and numerous encounters with students and other teachers.

The underlying theme, which the reader should keep constantly in mind, is that so long as we remain unaware of the extent to which our hidden culture determines how we practice, think about, and do research on teaching, attempts at reform are likely to be ineffective and we will remain locked in a system that inevitably produces failure and social inequality (Cuban, 1998).

THE JOURNEY BEGINS WITH A TAPE-RECORDER (1960–1968)

My research journey began when I was a graduate student interested in understanding how teachers shaped student learning. I persuaded a group of experienced teachers to let me record in their classrooms by hanging microphones from string attached to the light fittings. Analyzing 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 teachers were being very sensitive to the individual needs and interests of their students, the interaction took place within pre-

dictable structures and rules of social interaction (Nuthall & Lawrence, 1965). The following extract of an interaction recorded in a fifth-grade classroom in 1959 (it could equally well have been recorded in 1999) illustrates a typical pattern. The class was discussing a quotation from a poem: “With fingers weary and worn, with eyelids heavy and red, a woman sits in unwomanly rags, plying her needle and thread”:

Teacher: I wonder if anyone can tell me what time in history that is likely to be? Suzy?

Suzy: About the French Revolution. They're poor people.

Teacher: Very good answer. I wonder if you could give me a further reason. Peter?

Peter: Round about the Middle Ages?

Teacher: It could be. Len?

Len: I thought it might be about the time of the Bayeux tapestry. Doing tapestry and that.

Teacher: Yes, it could. But I think we have had an answer so far which was much closer to the real time.

Like all such whole-class discussions, this one was initiated by the teacher asking a question. A student responded and the teacher commented on the student's answer before asking a further question. Again a student responded and the teacher commented briefly (“It could be”) before inviting a further student to respond. Finally the teacher made a more general comment intended to guide the discussion in a specific direction. At the core of this pattern of interaction is the question-answer-comment sequence. Typically this pattern is repeated over and over with variations (more student answers, different kinds of comments) within the context of information provided by the teacher or other resources (in this case the quotation from a poem).

Later, when I worked as an assistant with a research team at the University of Illinois, I found that the same question-answer-comment patterns were being identified by researchers working in classrooms in different countries and in different European languages (Simon & Boyer, 1968). Stenographic records of teaching made in the early 1900s showed that these patterns already had a very long history (Hoetker & Ahlbrand, 1969). Teaching, like language, has its own underlying structural rules and, like language, it is simultaneously both highly creative and tightly rule-bound. If culture consists of those habitual action sequences that pattern social interaction (McLaren, 1993; Quantz & Magolda, 1997), then, although we did

not think this way at the time, we had begun to uncover the ritualized routines that make up the culture of teaching.

LEARNING THAT TEACHING EXPERIENCE MAKES NO DIFFERENCE (1968–1974)

The next stage in the journey occurred when I worked with colleagues at Christchurch College of Education to work out ways of incorporating understanding of the underlying patterns of teaching into the training of preservice teachers. We tried several ideas, including training a group of preservice teachers to analyze recordings of their own teaching. Then, working with Cliff Wright, we evaluated whether this training made a difference to the effectiveness of the way they taught. We compared the teaching of the group of preservice teachers who had been trained to analyze recordings of their own teaching, a matched sample of preservice teachers who had not had the training, and a group of experienced teachers who were experts in science teaching. Each of these preservice teachers and teachers taught the same material on the life of the black-backed seagull over the same length of time, but in their own way.

We made audio recordings of each teacher interacting with her students and assessed student learning using specially designed tests of the content of the lessons. We then correlated the characteristics of the teacher-student interactions with measures of the students' learning, corrected for prior knowledge and intelligence. The results revealed a pattern of teacher questioning and feedback that was significantly related to student learning (Wright & Nuthall, 1970). According to the research criteria of the time, this was a well-designed study, and it was reprinted as a model of process-product research (cf. Morrison & McIntyre, 1973; Skowronek & Schmied, 1977).

To our surprise, however, there were no discernable differences between the experienced teachers, the beginning teachers we had trained in lesson analysis, and the untrained beginning teachers in either 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 may be carried out in much the same way, with much the same effects, by novices and experts alike.²

THE JOURNEY SPLITS IN TWO: THE EXPERIMENTAL STUDIES (1974–1980)

At this point, my journey took two different paths that later joined up again. In one branch of the journey, I worked with two graduate students to

design and carry out experimental studies of the effects of teaching on student learning. In the other branch of the journey, I began to write a book for teachers on what the new classroom research said about how to teach.

Experimental studies seemed the way to go. Correlational studies of the kind that Cliff Wright and I had carried out deal only with the surface features of teaching. You cannot tell whether the behaviors that appear to be related to student learning are merely symptoms of other more significant aspects of teaching that have been overlooked or remain hidden from the researchers.

The two graduate students (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 interaction patterns while still responding spontaneously within those patterns. For example, they used a predetermined sequence of questions and responded using a predetermined set of alternative comments depending on the nature of students' answers (learned by heart by the researchers somewhat like the script of a play). Because the students were thoroughly familiar with the ritualized routines of whole-class teaching, recordings of this scripted teaching were indistinguishable from recordings of the everyday teaching of experienced teachers. The students did not realize the teacher talk was scripted but got annoyed when their standard attempts to sidetrack the teacher failed.

These were very successful sets of experiments. Church's experiments developed our understanding of the interactive effects on student learning of using different types of questions and feedback (Nuthall & Church, 1972). Hughes's (1973) experiments uncovered the effects of different ways of managing student participation. In my view, these studies remain some of the most detailed and well-designed experimental studies of teaching that have ever been carried out.

But these experiments uncovered major problems with this type of research. Because of their very precise and detailed design, they highlighted the enormous interactive complexity of teaching. Designing further experiments with the precision and control needed to produce clear results would have meant carrying out hundreds of large-scale experimental studies comparing the effects of the multiple ways teachers interact with their students. For example, Church taught in 42 different fifth-grade classes to get reliable data on the interactive effects of four different types of teacher questioning and feedback.

I now realized that so much of teaching is contingent, with an almost infinite variety of permutations and combinations of behaviors dependent on the multiple contexts in which it occurs. This kind of research would produce an enormous compendium of experimentally validated dos and don'ts. If teachers could digest such a compendium, it would turn them into

encyclopedic robots, with little or no understanding of why they were doing what they were doing. The experimental methods might satisfy the cultural rituals of the research community but would produce little of value, and much probable harm, to the professional life of teachers (Nuthall, 1989).

I began to think that instead of the expectations of the research community predetermining the kind of results the research should produce, we should first find out what kind of knowledge would be most useful for informing teachers' thinking and guiding their practice. Once that was understood, we could then design research methods that were specifically focused on producing that knowledge.

THE OTHER BRANCH OF THE JOURNEY: A BOOK ABOUT RESEARCH ON TEACHING (1974–1980)

On the other branch of my journey, I looked for every study I could find that could throw light on the way teaching shaped student learning. The purpose was to write a book that would make the best research information available to teachers. What I found, however, was a field in which the findings were extremely varied and often contradictory. Even among the studies that involved direct observation or recording in the classroom, there were surprisingly few replicated findings (Nuthall, 1970).

As I puzzled about the kind of research needed to produce knowledge that was useful to teachers, I reached the conclusion that at the very least it must produce replicable results (Fodor, 1968; Toulmin, 1961). If research is to have credibility, a teacher should be able to trust, with some certainty, that the results will hold up in different contexts. It would be unethical to expect a teacher to use a new method if the results of research on that method had never been independently replicated. It would be similarly unethical to expect a teacher to believe a theoretical explanation of the teaching process if that explanation had been based on the results of a single, unreplicated study.

At that time I knew, or had contact with, many of those 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, by examining the raw data for myself, whether the results of different studies were comparable. Examining this data was time-consuming and depressing. I found little relationship between the statistical results (e.g., the statistical significance of a correlation) and the patterns of data within different studies. For example, several studies used Flanders's (1970) system of interaction analysis and found significant negative correlations between student achievement and the teacher's use of criticism. However, the raw data showed that the teachers using the least criticism in one study were actually using more

criticism than the teachers using the most criticism in other studies. If a teacher were to try to imitate the behavior of the “good” teachers in one study, they would be using the same amount of criticism as the “bad” teachers in another study. We forget that statistical significance depends on the relative frequencies within a study, and not on the real or actual frequencies that have practical significance.

After a lot of detailed work on this original data, 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 different studies (including my own) that had apparently identical statistical outcomes.

THE JOURNEY REACHES ITS LOWEST POINT: SEARCHING FOR THE RIGHT METHODS

This was the lowest point in the journey, so far. If ever there was a time in the wilderness, this was it. It no longer seemed possible that research on teaching could ever produce reliable and useful results. What is more, the kind of painstaking research that I had done was now being publicly attacked for being mindless dust bowl empiricism, lacking significant theory, even fascist in intent (Bolster, 1983; Delamont & Hamilton, 1976; Garrison & Macmillan, 1984).

I was, however, more than sympathetic to one of these major criticisms. 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 (Tom, 1985). But acknowledging this did little more than redefine the problem. It was still not clear how we could include the subjective and the personal but still produce results that were reliable and replicable, results that teachers could trust and would form a solid basis for understanding the relationship between teaching and learning.

Meanwhile, the journals continued to publish studies that showed statistically significant results. Books were published that listed these results, not identifying, or glossing over, inconsistencies as the errors we ought to expect in such a messy and complex area (Bennett, 1976; Dunkin & Biddle, 1974; Peterson & Walberg, 1979). There were no systematic programs of research in which researchers added to each other’s findings. The problems of replication never became apparent and disappeared from the research agenda. It was generally believed that inconsistencies in data were due to random errors that canceled each other out (Brophy & Good, 1986).

THE JOURNEY TAKES A NEW TURN: STUDYING STUDENT LEARNING (1978–1984)

A way out of this wilderness came through working with a graduate student who was also an experienced and capable teacher. Adrienne Alton-Lee 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 one student learned a particular concept when another student did not.

She designed her dissertation so that she could observe the experiences of individual students and determine how those experiences translated into learning (Alton-Lee, 1984). This was an entirely new orientation for me. It required looking at teaching through the lens of individual student experiences. This in turn required a new way of recording and analyzing classroom data. This led to the invention of what we called the “item-file.” The purpose of the item-file was to collate the observations and records, both objective and subjective, of every experience an individual student had that related in any way to the learning of a single concept (assessed through “item[s]” in the achievement test; Gage, 1989). Each item-file was, as far as possible, the complete life history of a single concept in the mind and experience of a single student.³ Because these item-files contained every detail of each student’s experiences (from recordings, observations, and the student’s own point of view), they gave a sense of being as close to the real lives of students as it was possible to get.

Although Alton-Lee’s dissertation was, in many ways, a brilliant success (Nuthall & Alton-Lee, 1990), it was based on the observations (over 8 weeks) of just 3 students in one classroom. We needed to know if the findings could be replicated in other classrooms with other students. So the journey, for the next 8 to 10 years, was spent carrying out three increasingly sophisticated replications of the original design. Our technician (Roger Corbett) designed individual miniature microphones for each student to wear. We found a way of attaching sets of small video cameras (with zoom and wide-angle lenses) to the classroom ceiling to supplement live observations. We could now record in detail the experiences of individual students continuously over long periods of time through different kinds of classroom activities.

DISCOVERING THE SEPARATE WORLDS OF STUDENTS AND TEACHERS

The first thing that became apparent from these very detailed recordings and observations was how little teachers knew about what was going on in their classrooms. Even live observers keeping continuous written records of

the behaviors of individual students missed up to 40% of what was recorded on the students' individual microphones and the video cameras (Alton-Lee, Nuthall, & Patrick, 1993).

We were discovering the ways students live in a personal and social world of their own in the classroom. They whispered to each other and passed notes. They spread rumors about girlfriends and boyfriends, they constantly commented on each other's and the teacher's behavior, and they continued arguments that started outside school. It became clear that the students cared more about their peers' judgments than they cared about the teacher's opinion. Within this pervasive (but hidden) peer culture, sexism and racism were alive and flourishing even when the teacher actively promoted fully inclusive learning activities or believed she was treating girls and boys equally (Alton-Lee, Nuthall, & Patrick, 1987; Alton-Lee, Densem, & Nuthall, 1991).

The data also revealed the extreme individuality of student learning. Typically, students already knew at least 40% of what the teachers intended them to learn. Consequently the students spent a lot of time in activities involving what they already knew and could do. This prior knowledge was, however, specific to individual students, so the teacher could not assume that more than a small fraction was common to the class as a whole. Because of these individual differences in prior knowledge and in the way students engaged with classroom activities, a third of what each student learned was not learned by any other student in the class, and the rest was learned by no more than three or four others (Nuthall, 1999a).

It was about this stage in my journey that I began to understand the significance of the difference between the teacher's perception of the classroom and the students' perceptions. In order to take responsibility for a class of 25 to 35 students who have different knowledge, skills, interests, and motivations, teachers have to develop ways of managing the class as a whole. It is impossible to focus on the individual learning of any one student for more than very brief periods.

The ritualized routines of teacher-student interaction appear to have evolved to solve this problem. Teachers follow patterned action sequences (Bohannon, 1995) that work because both teachers and students have learned exactly what to expect of each other (Quantz & Magolda, 1997). Within these ritualized patterns, students learn how to manage 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 the teacher is keeping more than a passing eye on them, as the teacher will get upset if the students do not respond in culturally expected ways (Hughes, 1973).

I began to think of the class as an orchestra following a cultural musical score. So long as everyone knows his part, the whole works together

effectively. Furthermore, when everyone is following her own score, the coherent sound of the whole orchestra makes it very difficult to separate out the sounds of individual instruments. In a classroom, we have found that even a trained observer finds it extraordinarily hard to focus on the continuous experience of an individual student without getting absorbed back into the performance of the class as a whole (Alton-Lee et al., 1993).

IDENTIFYING HOW STUDENTS LEARN FROM CLASSROOM EXPERIENCE (1990–1995)

The next significant stage in my journey occurred during 1990 when I took study leave to work full-time on analyzing the data from the three studies. My primary focus was on the way classroom experiences affected student learning. For 9 months of that study leave, I did nothing else but classify the teacher's and students' behaviors and correlate them with outcome measures. I used everything I knew about formal and informal data analysis, producing endless printouts of tables and graphs. But gradually it dawned on me that for all the richness of the data we had obtained, the old problem still remained. Alton-Lee's dissertation results were not replicated. Each of the new studies showed different relationships between teacher and student behaviors and achievement even when nested within different classroom organization and activity structures.

I kept thinking that if I added more variables, included more detail, used more complex analysis, patterns would emerge. They did, but they didn't replicate across the three classrooms. At this point in the journey, it began to seem that unless we could get inside the minds of students, we would never understand exactly how learning occurred and how teachers shaped that learning.

However, I decided, as a last resort, to take one student's learning of one concept and try, from our detailed recordings and observations, to get an insight into what was going on in that student's mind. It was an 11-year-old boy (code-named Jon) and it was an item-file about recent migration to New York from within a social studies unit.

As I examined the sequence of his experiences, the things he saw, heard, read, said, wrote, and joked about, as he learned that Puerto Rico was a major source of migrant labor, the obvious dawned on me. The most significant thing about learning is that it involves 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. This meant that the same experience might have quite different effects depending on where it occurred in the sequence of an individual student's experiences. The methods of data analysis that I and others were using in this

research assumed that teacher and student behavior variables had stable effects regardless of when or where they occurred. Our methods were not sensitive to the possibility of constantly changing effects.⁴

It also dawned on me that the most important aspect of a student's experiences, from the point of view of conceptual learning, is the information (or meaning) that the student extracts from those experiences. Students cannot learn from information they are not exposed to, regardless of the kinds of questions the teacher is asking, the kinds of resources the student is exposed to, or the context in which the learning is taking place.

In hindsight, all this seems very obvious. But like most other researchers, I was caught up in the culture of the research community with its focus on general variables of theoretical significance across many different physical and temporal contexts. I had been blind to the significance of understanding the particular.

During the remainder of that year, I developed ways of classifying the content, timing, and sequencing of the information that the students extracted from their experiences. My analysis of Jon's experience showed that when he learned a concept, he had experienced, on at least three separate occasions, all of the information he needed to fully understand the concept. Each full set of information could come from a single experience or could be made up of different kinds of information coming from several different experiences. His mind was presumably actively collating and integrating these different kinds of information.

LEARNING TO PREDICT PRECISELY WHAT STUDENTS WILL AND WILL NOT LEARN

On the basis of the analysis of Jon's experiences, I developed a procedure for analyzing the content of the information that students encountered and identifying when a student had encountered the three complete sets of the information that seemed to be needed to understand, learn, and remember a concept. It is important to note here that it is not simply repetition of information that is involved. Most of the time the information that students encounter that is relevant to specific concepts is partial or fragmentary, involves different analogies or examples, and occurs in different contexts and forms. What is significant is that it adds up to the equivalent of three sets of the complete information needed to understand a concept. The procedures I developed were largely concerned with determining how partial information, analogies, examples, and other related pieces of information added up to a complete set of information. Once this was worked out, I applied the procedure to the experiences of another 3 students in Jon's class, made some minor adaptations, and then used it to predict ex-

actly which concepts the students in the other two studies would learn or not learn (for details of the procedures see Nuthall, 1999a; Nuthall & Alton Lee, 1993).

These procedures worked. They successfully predicted, with 80–85% success, the learning of nearly 500 concepts through the experiences of 11 students in three different classrooms. It involved a lot of detailed analysis of individual student experience (what has now become known as “micro-genetic” analysis; Siegler & Crowley, 1991), but it was replication of the kind I had never seen before (Nuthall, 1999a; Nuthall & Alton-Lee, 1993).

Let me deviate from the narrative at this point to give an example of how these procedures worked. The example is taken from one of our later studies in which a Grade 7 class was studying Antarctica as an integrated science and social studies topic. One of the apparently simpler pieces of information that the teacher wanted the students to learn was that Antarctica is the driest continent in the world. It follows the parallel experiences of three students: Joy, Teine, and Paul.

On the 2nd day of the unit, the teacher showed the class a video on Antarctica. At one point in the video, the camera showed extensive scenes of ice and snow, and the commentator said (among other things):

Another surprise is that it doesn't snow as much as you'd expect, except around the coast. Almost no snow falls on the high central plateau. In fact less moisture falls here each year than in the Sahara Desert, making Antarctica overall the driest continent. The little that does fall can't melt so it accumulates bit by bit. Over millions of years it's compacted into a vast sheet of ice covering virtually the whole continent.

Our observations and video recordings showed that Joy and Paul appeared to be watching and listening to the video. Paul made occasional notes about temperatures, wind speed, and the ice cover, but not about rainfall. Teine was engaged in passing notes to her friends. Before the video started, Teine and her neighbor Leigh had a whispered discussion about boyfriends, part of which went as follows:

Leigh (whispering): Just try to be mature about this.

Teine (also whispering behind her hand): Yeah I know. But if she wants John—I mean she never acts as if she wants him. You know, like, I was . . .

During the video, this interaction was continued by passing notes to each other. For example, the following is a summary of the observer's notes on Teine's behavior over the first 4 min of the video:

Reads note from Abbie passed by Leigh. Scratches head and writes on note. Passes note to Abbie. Glances at video and looks at Abbie. Leans down to pick something up off floor. . . . Receives note from Abbie. Watches video and puts note aside. . . . Sighs, rubs eyes and glances at video. Takes Abbie's note, rubs eyes, glances at video while fiddling with Abbie's note. . . . Writes on the note and passes note to Abbie. Looks across at Leigh and Abbie. Watches video.

Immediately following the video, the teacher held a whole-class discussion about the content of the video.

Teacher: OK. Put your hand up if you've learnt one thing from that [video] and you might like to share it with us. OK. Chris?

Chris: It's um, it's a dry place.

Teacher: It's one of the driest places in the world. Why would it be a dry place do you think?

Colin: (inaudible)

Teacher: Because ice isn't wet did you say?

Colin: Yeah, well when it (inaudible)

Teacher: Right. OK. That's something that surprised me, that it's a dry place. Good.

About 10 min later in the same discussion, Joy's close friend Jane made the same point.

Jane: It's even drier than the Sahara.

Teacher: OK. That the, that Antarctica is drier than the Sahara Desert. That might be a really interesting thing to find out. Why is it drier than the Sahara Desert? OK.

These statements repeat some of the information contained in the video commentary: that Antarctica is dry, that it is drier than the Sahara Desert. The teacher also repeated the claim in the video that this is surprising and really interesting. Joy listened and contributed to this discussion. Neither Paul nor Teine contributed, but they appeared to be listening.

The following day the pupils were working in groups studying photographs of people working in Antarctica. Their task was to identify what the people in the photograph were doing and why. The following discussion developed in the group that contained Joy and Paul (Teine was in another

group). Their photograph was of a person looking at some equipment on a tripod on the snow.

Maude (speaking of the person in the photograph): . . . Studying weather, yeah.

Paul: But this could be rainfall, rainfall for the week.

Koa: Huh?

Paul: Could be rainfall for the week.

Koa: Could be.

Maude: How do you know rain falls?

Joy: I didn't know rain falls in Antarctica.

Paul: Rain does fall in Antarctica.

Maude: Amazing!

Paul: That's what turns this into ice.

Maude: I thought, I thought, I thought it was dry.

Paul: So it still has a bit of rain.

It appears that Joy had retained in her working memory what she heard in the video and the class discussion. Paul, who also probably remembered the class discussion, still maintained his prior belief that ice implies rainfall. In Teine's group, studying another photograph, there was no reference to the dryness or lack of rainfall.

Two days later the teacher led a whole-class discussion of the differences between living in the city where the students lived and working in Antarctica:

Teacher: . . . What would you not see? . . .

Nevin: I was just going to say, you wouldn't see rain.

Teacher: You wouldn't see rain. Why not?

Nevin: 'Cause it's the driest place in the world.

This time another student, Nevin, stated that there is no rain in Antarctica and gave the explanation that it is because Antarctica is the driest continent. Joy, Teine, and Paul appear to have been listening to Nevin's statements.

The next day, the students were asked to write a report about what they had learned so far about Antarctica. Joy wrote in her report that she had

learned that “It hardly ever rains in Antarctica. . . . Antarctica is the driest country.” Neither Paul nor Teine made any reference in their reports to dryness or lack of rain.

What the teacher wanted the pupils to learn was that Antarctica is the driest of all the continents. As Table 1 shows, Joy encountered this statement on three occasions (1, 2, and 4 in the table). She identified it as something she had learned in her report (5). She seems to have understood this to mean that it did not rain in Antarctica and confronted Paul with this fact in their group discussion (3). However, by the time she wrote her report she had changed it to “it hardly ever rains.”

The three direct encounters with the relevant information (1, 2, and 4 in the table) and the additional encounter when she wrote her report (5) ensured that Joy remembered this information. When we asked her 12 months later what she remembered about the climate in Antarctica she said:

It [rain] almost never falls. . . . It’s um . . . it’s too cold for it or something I think it is. It freezes up into snow or something.

When we asked Paul about rain, he replied:

Paul: Almost never falls. Because it, oh the rain almost never falls but there’s snowstorms. Stuff like that.

Interviewer: Did you know that beforehand or did that come up during the unit?

Paul: No. It came up during the unit.

Table 1. Summary of the experiences of the 3 students related to learning the idea that Antarctica is the driest of all the continents

<i>Experiences of relevant content</i>	<i>Joy</i>	<i>Teine</i>	<i>Paul</i>
1. Video: “less moisture falls than in the Sahara . . . the driest continent”	Attending	Passing notes	Attending
2. Class discussion: “one of the driest places in the world . . . it’s even drier than the Sahara”	Attending	Attending	Attending
3. Peer group: (Joy) “I didn’t know rain falls.” (Maude) “I thought it was dry.” (Paul) “still has a bit of rain”	Involved in group discussion	In another group	Involved in group discussion
4. Class discussion: “you wouldn’t see rain . . . it’s the driest place in the world”	Attending	Attending	Attending
5. Writing a report: (Joy) “it hardly ever rains . . . is the driest country”	Writing report	Not in report	Not in report

Interviewer: Think for a minute and see if you can remember in what way it came up.

Paul: Um, I think it was when they were telling us about it being the driest continent and that it hardly ever falls.

Interviewer: Yeah? And was that a person or a video or?

Paul: Um, video.

Paul had four relevant experiences (1, 2, 3, and 4 in Table 1). This was sufficient for him to recall the basic information (“the driest continent”) 12 months later but was not enough to ensure a clear understanding. He had forgotten that the voice on the video said it hardly ever snows. This was mentioned only once and so was not remembered, leaving him with his prior belief that it snows a lot in Antarctica alongside his newly acquired belief that it is the driest continent.

Teine’s experiences were different. She probably did not hear much of what was said on the video. Her group did not discuss anything relevant to the climate in Antarctica (3 in Table 1), and she did not include any reference to climate in her report of what she had learned. As a consequence of less than three encounters with the relevant information, when Teine was asked in her interview about rainfall, she did not remember anything about Antarctica being the driest continent or about the lack of rainfall.

The predictive success of these procedures for analyzing the content of student experiences suggested that understanding the learning process was simpler than I anticipated. It depended primarily on having detailed information about the continuous experiences of individual students and of being able to collate the experiences according to their relevance to individual concepts. It also suggested that we needed to make a careful distinction between short-term and long-term memory. In the above example, Joy recalled the content of the video and the following class discussion when she participated in the discussion with her group. She also recalled this information when writing her report. This recall was a function of her short-term working memory, which, we discovered appears to store information for about 2 days (Nuthall & Alton-Lee, 1993). There is no reason to believe that Teine did not also store the content of the class discussion in her short-term working memory. The difference is that the accumulation of at least three different sets of relevant information in Joy’s short-term working memory resulted in the cognitive processing of that information (such as integration with previous knowledge, identification of implications, and evaluation of consistency with related knowledge; see Nuthall, 1999b) that transferred the newly constructed concept to her long-term memory. For

Teine, that lack of sufficient sets of information in short-term working memory resulted in that information's being forgotten.

Understanding the learning process in this way seemed like a significant discovery. Although there was nothing like it in anybody else's research, it felt as if the journey was headed in the right direction. It also suggested why this key to learning had remained hidden. It is commonly assumed that if a teacher explains a new concept clearly and vividly, the explanation will become part of a student's knowledge (see, for example, explanations of learning in commonly used textbooks such as Gagné, Yekovich, & Yekovich, 1993). But our data indicated that experiences like this do not work individually. Unsupported by integration with other relevant experiences, the student will rapidly forget specific experiences. Since many of a student's experiences are hidden from the teacher, there is no way the teacher can know how often the student engages with relevant content.

With this kind of detailed analysis of the learning process it became clearer what role the teacher played and how the learning process was affected by interactions with peers and self-created learning experiences.

DISCOVERING HOW STUDENTS USE THEIR KNOWLEDGE TO ANSWER QUESTIONS ON TESTS

Another significant discovery occurred at this time. Critical to our studies had been the very detailed ways in which we assessed individual student learning. We used carefully developed and administered tests but always supplemented them with extensive individual interviews that explored each student's knowledge and understanding in greater depth. We also spent time analyzing how the students used their memory for classroom experiences to answer test questions and explain their knowledge and learning during interviews (Nuthall & Alton-Lee, 1995).

Out of this came the insight that testing is like interviewing. Both depend on the relationship between tester and student. How the student responds depends on what the student thinks the tester or interviewer wants and will do with the responses. 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 set of relationships between the tester, the student, and the purposes of the assessment.

I also came to understand that what a student knows is a coherent body of beliefs and understandings that is for the most part logically and consistently interconnected. Distinctions among facts, concepts, principles, generalizations, and procedures exist more in the theories of researchers than they do in the minds of the students. Knowledge is an almost seamless network of complex relationships without finite boundaries rather than the

set of independent countable objects that is the implicit basis for standard measurement theory (Wainer & Braun, 1988).

This led to the conclusion that the scores that students get on standard paper-and-pencil tests do not necessarily represent what the student knows or can do but are determined initially by the students' interests, motivations, and understanding of the purposes of the test.

To get a sense of this, imagine a classroom on a very warm afternoon. A student (let's call her Amy) distractedly watches a teacher she doesn't like hand out printed tests and say something about how important these tests are. When the teacher tells the students to begin, Amy answers a few easy questions at the beginning. But as the room gets warmer, her clothes get stickier, and the questions get harder, she wonders why the hell she needs to do this. She starts doodling patterns down the side of the answer sheet as the thought crosses her mind that no one she knows gives a damn about what she does on the test. She sucks her water bottle, doodles a few more faces on the answer sheet to keep boredom at bay, and waits for the test to be over and for real life to begin again.

Amy was participating in an international study of mathematics achievement (Nuthall, 1978). There is no reason to believe that taking a test is not like other classroom activities that Doyle has described as a form of "work" (Doyle, 1986). Some students will attempt to do their best because they want to sustain their higher status with respect to their rivals in class or because they are anxious about their parents' possible reaction to a low score. Other students will put in little effort because, like Amy, they see little point in the test or want to sustain their reputation as antischool rebels. It is naïve to see test results as somehow independent of students' interests, motivations, and beliefs about the values and purposes of tests.⁵ Understanding this throws considerable doubt on the validity of most research on teacher and school effectiveness (e.g., Sammons, Hillman, & Mortimer, 1995) and on standards of student achievement (especially the large international studies such as TIMSS and the Organization for Economic Cooperation and Development's Programme for International Student Assessment) that depend on paper-and-pencil tests and test administration procedures that have no personal significance for students. Important as this issue seems, I have never been able to find research that looks at the relationship between student test-taking motivation and achievement test scores. The closest research shows that the scores of job applicants on cognitive-ability tests are strongly affected by their need for the job and their beliefs about the validity of the test and its role in the selection process (Chan, Schmitt, Sacco, & DeShon, 1998; Sanchez, Truxillo, & Bauer, 2000). Again, this insight resulted from abandoning the view from above and adopting the view from below, seeing testing through the eyes of the student and not the teacher or administrator.

THE JOURNEY TAKES A DRAMATIC TURN: QUESTIONING THE
ROLE OF ABILITY (1995–1998)

These discoveries were, however, only preliminary to the journey's taking a more exciting and disturbing turn. Alton-Lee and I carried out two further studies that allowed us to develop and extend the prediction procedures to cover further kinds of learning and to test their predictive validity with older students learning different kinds of concepts. Again the prediction procedures worked just as well. They were now objective and formal enough to be computerized. The tally was now the successful prediction of 1,100 different concepts, learned or not learned, by 21 students in five different classrooms (Nuthall, 1999a, 1999b).

However, when I was asked what role student ability or intelligence played in our predictions, I had to answer that, as far as our data were concerned, none at all. Given the same experiences, the less able students appeared to learn from their experiences in exactly the same way as the more able students. In the example given above of the way Joy, Paul, and Teine learned or failed to learn that Antarctica is the driest continent, their relative levels of ability were not relevant. All that was needed to predict their learning was an analysis of the content of the relevant information they encountered. But this finding contradicted all the research that shows that academic ability or intelligence is the best single predictor of school achievement. Was the view from below so very different from the view from above (cf. Gagné & St. Pere, 2002)?

Our data did show that the more able students started with more background knowledge and ended up learning more than the less able students. So the question was: What was different about the experiences of the more able students that gave them this advantage?

Analyzing the differences between individual students' experiences (such as the differences between Joy's and Teine's experiences; see Table 1) led to the discovery that a significant proportion of each student's learning experiences was either self-selected or self-generated, even in quite traditional classrooms. Because we were able, from the prediction procedures, to isolate the specific experiences that were critical to the learning of individual concepts, we discovered that the more able students generated more critical learning experiences for themselves (see, for example, the differences between Joy's and Teine's experiences above). The more able students talked more with other students about relevant content. They asked more questions and seemed more interested, more persistent, and less likely to be distracted. Table 2 shows the proportion of concepts learned during the Antarctica unit by Paul, Joy, and Teine in which the critical learning experience (without which the concept would not have been learned) was due to the teacher directly (such as a teacher-managed class discussion), the

teacher indirectly (through teacher-designed activities), the student's own activities (such as personal use of resources), or spontaneous talk between students. Concepts for which there were significantly more learning experiences than were needed to learn and remember the content were not included. Note that in order to provide a context, the table also contains the actual number of concepts that the teacher intended the students to learn that were already known, learned, and not learned by each of the students during the unit on Antarctica.

Paul, the highest-ability student, depended least on teacher-directed activities and most on self-created activities and spontaneous peer talk. Teine, the least-able student, depended most on teacher-managed activities. As the example of learning about Antarctica's climate shows, much of the spontaneous talk between Teine and her peers was about noncurriculum issues. The following excerpt, taken from another of our studies, illustrates the way a spontaneous conversation among interested students can contribute to the content the students experience. A Grade 7 class was studying ancient Egypt. Adam, Dean, Judith, and Kenneth were working as a group looking through resource books for evidence of common occupations in ancient Egypt. Adam was reading about the life of the Pharaohs:

Adam: Did you know when a pharaoh got buried there was always a war over who would be the next pharaoh?

Dean: Why?

Table 2. Percentage of learned concepts for Paul, Joy, and Teine for which a critical learning experience was teacher-managed, self-selected, or self-created

	<i>Paul</i>	<i>Joy</i>	<i>Teine</i>
Academic ability (percentiles) ^a	89	70	34
Number of concepts			
Already known	120	109	90
Learned during unit	72	61	41
Not learned during unit	31	55	89
Percentage of learned concepts in which a critical learning experience was during			
Teacher-managed activities	39.3	48.9	64.5
Choice of teacher-designed activities	32.8	26.7	22.6
Self-designed activities and use of resources	13.1	8.9	6.5
Spontaneous peer talk	14.8	15.6	6.5

Source: Adapted from Nuthall (2001a).

^aAverage percentile for age score on nationally standardized tests of reading comprehension, reading vocabulary, listening comprehension, and mathematics. For details, see Nuthall (1999b).

Judith: 'Cause they didn't have a baby.

Adam: Wonder what they done in the hospitals when they have a baby.

Dean: Same thing we do here.

Judith: They didn't have hospitals.

Kenneth: The only thing that they had, they had midwives.

Dean: Yeah I know.

Judith: Oh yeah, the midwives.

The talk changed to a discussion of embalming, until Judith brought the discussion back to birth:

Judith: They would have to go through all the pain, though, eh?

Dean: What?

Judith: When they had the baby.

Kenneth: What? They don't have, like, sleeping pills back then.

In this exchange, Kenneth introduced the other students to another occupation (midwife) that might have existed in ancient Egypt. This kind of spontaneous exchange depended on the interests and background knowledge of students but was more likely to occur among higher-achieving students. There was no evidence that the higher-achieving students found the tasks easier or that their minds processed their experiences differently (Nuthall, 1999b, 2001b).

It was Piaget's work that provided the metaphor that helped me understand this puzzle. 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 processes and extracts what it needs from experience.

Within reasonable limits, the learning processes, like the metabolic processes that take place in the body, are universal across our species. How the metabolic processes work in an individual depends on the kinds of food the individual eats and the individual's energy output. In the same way, how the learning process works in an individual depends on the kinds of experiences the individual has and the ways the individual uses those experiences.

This means that those students who use the classroom and its activities to further their own interests and purposes learn more than those who dutifully do what they are told but do not want, or know how, to create their own opportunities. If background interests and prior knowledge determine

what students learn, then differences in apparent academic ability are just as likely to be the product of differences in classroom experiences as the other way round.⁶ This led me to think of the relationships among ability, learning, and classroom experiences as mutually interactive within a context of the students' social roles and relationships.

THE JOURNEY AGAIN TAKES TWO DIRECTIONS: LEARNING FROM MY STUDENTS (1990–2000)

At this point, my journey again split into two parallel paths. Along one path, I set out to discover why learning a new concept or changing a student's belief required the sequence of experiences that our research had identified. What, I wondered, goes on in the student's working memory as each new experience is processed and related to subsequent experiences? If Vygotsky is right, then the way students process experience in school is the result of internalizing the social structures and routines of classroom life. We had already identified the recurring patterns that made up whole-class activities, but we needed to discover the recurring patterns that structured small-group and individual activities (Nuthall, 1999b, 2000a).

The other path of the journey involved my experiences teaching teachers. For many years I taught a course on research on teaching for experienced teachers and teacher educators. Two assignments in that course provided significant insights into how these teachers and teacher educators thought about teaching. In one assignment, I asked the course participants to interview fellow teachers about how they knew when their teaching was effective. Following a method developed by Jackson (1968), their fellow teachers were asked to describe the classroom evidence they used to judge the effectiveness of specific teaching incidents.

This assignment always produced the same results as Jackson's original study. Teachers know their teaching is going well from signs that their students are actively engaged with learning activities. They monitor the look in their students' eyes, their enthusiasm, their puzzlement, the questions they ask. In most teachers' minds, the criteria for successful learning are the same as the criteria for successful classroom management (Jackson, 1968).

The second assignment I gave the students was much more difficult. I asked these experienced teachers to carry out research on their own teaching. They were to record a small teaching activity or procedure, measure (using pre- and posttests or interviews) 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, the majority of these teachers and teacher educators carried out the assignment in the same way. They recorded what they and their students did together and measured what the students learned through a test and/or interviews. They then related the measure of the students' learning to indicators of how engaged the students were during the teaching activity. And they almost always 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. I wanted more than a study of classroom management. I wanted them to connect the content of their teaching to what the students learned. This requires identifying exactly how a student's knowledge or skills have changed and what specifically caused that change to occur. Knowing that a student is busily engaged in an activity does not tell you what (or how) the student is learning. You need to know exactly what information or knowledge is engaging the student's mind. To give a simplified example, it is not enough to say that a student learned because the student was busy reading a book unless you also identify what the student was reading and how that content related to what the student already knew.

What I learned from these experienced teachers 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 actually about learning (Fischler, 1994; Nunan & Lamb, 1996). 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 (Clark & Peterson, 1986; Leinhardt & Greeno, 1986; McNair, 1978).

As the 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 students were learning."

IDENTIFYING HOW STUDENTS AND TEACHERS MANAGE CLASSROOM ACTIVITIES (1998–2001)

On the other path of my journey at this time, I started a new series of studies using more extensive and sensitive methods of recording individual student experiences. The purpose was to get closer to students' minds and try to find out how the social structures and routines of the classroom shaped their experiences and their learning (Nuthall, 2001a). Guidance came from Piaget's and Vygotsky's theories of internalization and from recent research in cultural psychology (M. Cole, 1996) especially activity

theory and research on the cultural roots of classroom practices (Lave, 1991; Nuthall, 1997; Rogoff, 1993).

What emerged from these studies was a growing understanding of how classroom activities are structured and of how teachers and students create their own routines and patterns of interaction within the larger culturally determined patterns of classroom behavior.⁷ For example, the learning activities that teachers designed were usually made up of a sequence of four phases: an instruction phase, an activity phase, a reporting phase, and a wrap-up phase (Nuthall, 2001a). Within each of these phases, there were standard behavior routines and rules that structured the teacher's and students' expectations and interactions. During the instruction phase, for example, the students played the role of attentive listeners, paying just enough attention to answer any questions the teacher might ask them (Amerine & Bilmes, 1988). They were aware that only some students (mostly male) were ever asked questions and adjusted their attention accordingly. Where they could, they tried to negotiate the activity requirements downward using reasons they knew would be effective with the particular teacher providing the instruction (Doyle & Carter, 1984; Good, McCaslin, & Reys, 1992).

If the learning activity (e.g., a science experiment, a social studies research question) followed a small-group format, the established routines for social interaction came into play. The students followed their own rules about what counts as being helpful, contributing, copying, and cheating. Using their knowledge of which student usually knew what the teacher wanted, knew the answers, or knew best how to carry out the activity, they negotiated with that student to find the easiest way through the activity. They knew that, despite the teacher's instructions, the only significant outcome of the activity was completing the report or worksheet (Shanks, 1994).

Although the learning activity was supposed to produce learning, neither the teachers nor the students talked about learning. Instead, teachers talked about resources, about how long an activity should take, and what would happen if it was not finished on time (Nuthall, 2001b). Students' on-task talk was about the same things. They constantly compared how much each of them had completed. They asked how long the activity would take, whether the headings had to be underlined, where you could find the answer, and whether it had to be finished for homework (Nuthall, 2001a). When I showed students video recordings of themselves in class and asked what they were thinking during these activities, their most common response was that they were thinking about how to get finished quickly or how to get the answer with the least possible effort (Nuthall, 2001c). In short, the classroom life of teachers and students (outside of whole-class routines) is mostly about organizing and getting tasks completed. Doyle (1986) made the same point when he described teaching as being about (a) creating and sustaining order in the classroom and (b) moving students through the curriculum.

Detailed analyses of how learning activities were carried out in different classrooms revealed that every classroom was both unique and the same. Underneath the apparently unique sequence of events, there were predictable patterns and structures that coordinated the interactions of the participants. However, it also became increasingly apparent that the students were not just engaged in a single cultural context but were simultaneously involved in three distinct but interacting cultural contexts (Nuthall, 2001a). There was the visible cultural context of the public teacher-managed activity routines and rules. There was also the largely hidden but powerful cultural context of peer relationships and interactions. This peer culture extends across any setting in which students interact with each other and internalize the values and rules about roles and status appropriate to their age group (Benenson, Apostoleris, & Parnass, 1998; Ladd, 1999; Wentzel, 1999). And finally there was the personal cultural context of beliefs, attitudes, and knowledge that individual students brought into the classroom from their family and neighborhood backgrounds (Alton-Lee et al., 1993). It became apparent that understanding students' participation in classroom activities (and the consequent shaping of their minds through internalization) required an understanding of these three separate cultures and the ways in which students simultaneously affected, and were affected by, these cultures.

WHERE HAS THE JOURNEY TAKEN ME? UNDERSTANDING THE ROLE OF RITUALS

Let me pause here to review where the journey has taken me. Increasingly my research looked different from the standard ways of doing research and the findings (based on viewing the classroom through students' eyes) were distancing me from the standard ways of understanding teaching and learning. I was beginning to believe that much of the published research on teaching effectiveness, academic ability, and achievement was based on myths or misunderstandings rather than the realities of students' classroom experiences. However, the accumulating series of data-based insights seemed to be getting me closer to a valid understanding of how classroom teaching relates to individual student learning.

I had, by this time, come to understand that there are very good reasons why teaching and teacher-managed classroom activities followed predictable routines that were only hypothetically linked to student learning. Although the analogy with an orchestra helped to explain why teachers have so little awareness of the behaviors of individual students, it failed to explain why teachers seemed to have so little awareness of student learning. Unlike the conductor who can hear the players in the orchestra, the teacher is largely cut off from information about what individual students are learn-

ing. Teachers are forced to rely on secondary indicators such as the visible signs that students are motivated and interested. They are sustained, however, by the commonly held belief that if students are engaged most of the time in appropriate activities, some kind of learning will be taking place. For example, the class brainstorm is a typical ritualized classroom routine. It is carried out in the belief that a brainstorm is a good way to find out what students know about a topic and is used in the constructivist classroom as the basis for further activities (e.g., Roth, 2001). The brainstorm is a teacher-led whole-class routine that is initiated by the teacher's asking the students to share the ideas they have about a 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. Most teachers (and most observers) are convinced that if this routine is followed faithfully, it is an effective way of finding out what all the students know. In fact, teachers depend on the responses of a small number of key students as indicators and remain ignorant of what most of the class knows and understands (A. L. Cole, 1988; Dahllöf, 1967; Sahlstrom & Lindblad, 1998).

So long as ritualized routines, like the brainstorm, are recognized by teachers (and the society at large) as the right way to run classrooms and students have learned to play, and accept without question, the reciprocal roles that these routines require, then teachers do not have to pay more than passing attention to the experiences of individual students. For example, so long as the routines are following the expected format, well-intentioned teachers can be completely unaware that they are favoring boys over girls (Alton-Lee et al., 1991) or majority over minority students (Spindler, 1999) or teaching content that the students already know (Nuthall, 2001b).

Elbaz (1983) has made a similar point. On the basis of studies of teacher thinking, he argues that "the teacher's feelings, values, needs, and beliefs combine as she forms images of how teaching should be, and marshals experiences, theoretical knowledge and school folklore to give substance to these images" (p. 134). What I have come to believe is that the "feelings, values, needs, and beliefs" are the result of socializing into the commonly held culture of teaching and that the "theoretical knowledge and school folklore" are the result of the web of supporting myths and beliefs that underpin the cultural routines of teaching.

IDENTIFYING THE WEB OF SUPPORTING MYTHS: BELIEFS ABOUT ABILITY AND LEARNING

Within a culture, the ritualized routines that structure social interactions are unlikely to survive without a web of supporting beliefs or myths that explain and justify the way these routines are played out. As I became increasingly

conscious of contradictions between what I was finding out about students' classroom learning and commonly held beliefs about teaching and learning, I became aware of at least two sets of beliefs that worked to sustain faith in standard classroom routines.

Possibly the most significant is the set of beliefs we have about academic ability. When a teacher manages a classroom so that all the students are, as far as the teacher can see, busily engaged in appropriate activities, it is difficult to explain consistent differences in what the students produce in class or score on achievement tests. Teacher and students are carrying out the expected routines, so why do some students do well and others fail?

The concept of academic ability or intelligence explains this problem by transferring responsibility to the student rather than the teacher. Carroll's (1963) well-known model of school learning describes this set of beliefs perfectly. According to Carroll, learning is a function of time and ability. If a teacher can engage students in a learning activity and keep them "on task" for an appropriate length of time, then differences in achievement must be the result of differences in ability.

Our data, however, show that differences in what students learn, and differences in what they do on tests, are both created by differences in how they engage with classroom learning and testing activities. In both cases, their performance is a function of their interest, motivation, and understanding of the purposes and beliefs of the teacher and/or tester (Nuthall, 2001b). The result is that ability tests are highly correlated with school achievement tests (and with the kinds of work that students produce in class) not because of the role that ability is supposed to play in the learning process, but because scores on both ability and achievement tests are determined by interests and motivational and cultural factors. The high correlations between ability and achievement tests are, in a sense, a measure of the relative ineffectiveness of teachers at managing classroom learning.

Oakes (1985; Oakes, Wells, Jones, & Datnow, 1997) has detailed how the belief that ability determines learning becomes a self-fulfilling prophecy when ability-based tracking deprives low-scoring students of resources and lowers the quality and content of what they are taught. This set of beliefs also creates the sociocultural context in which students acquire, through interactions with each other, false but self-fulfilling beliefs about their own ability (Hallinan, 1996; Nuthall, 1999b).

The second set of beliefs relates to our understanding of how teaching is related to learning. The commonly held belief (which is a myth) is that school learning is a direct consequence of teaching. In the traditional view of teaching (expressed by the deputy principal in the quote above), it is assumed that whatever the teacher explains or demonstrates automatically and simultaneously becomes part of the knowledge of the student. It is this

view of learning that justifies the teacher's standing in front of the class and talking.

This myth is currently changing to accommodate new knowledge about students' misconceptions and conceptual change (Derry, 1996). It is now more widely understood that students learn by constructing their own knowledge (Ball & Bass, 2000; Hogan, Nastasi, & Pressley, 2000). But this constructivist view still contains most of the older myth. Student talk has replaced teacher talk, so that it is now assumed that when students talk together about a new concept or procedure they are "mutually constructing" knowledge (Wells, 2002). What has changed is the concept of what constitutes an effective "learning activity." What has not changed is the mythical belief that engaging in learning activities (such as listening to the teacher talking, discussing the results of an experiment, or writing a report of an investigation) transfers the content of the activity to the mind of the student. The concept of "time on task" as an explanation of learning, without regard for the content of the task, encapsulates the logic of this myth.

This myth deals with the problem that learning is invisible and cannot be seen in the activity of the teacher or student. It also obscures the fact that as learning occurs, so does forgetting. However, the data from our studies show that learning takes time and is not encapsulated in the visible here-and-now of classroom activities. Students are continuously engaged in simultaneously learning many different concepts that interact with each other. A student's network of knowledge is constantly changing as the information extracted from new experiences interacts with previous knowledge. Selective forgetting is also constantly changing the configuration of student knowledge (Nuthall, 2000b, 2000c). This makes it impossible for teachers to judge what their students are learning without much more detailed and individually differentiated data than they have available in the classroom.

WHAT SUSTAINS THE RITUALS AND THEIR SUPPORTING MYTHS?

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

First, there is a tendency to assume that teaching and face-to-face conversation are largely the same. Linguists describe them as different forms of discourse (Hicks, 1993). We know, through many years of participatory experience, the ritualized routines that produce satisfactory communication between persons in our culture. We know, usually without conscious aware-

ness, the techniques of telling and persuading. Consequently, it is easy to assume (and most of those outside the teaching profession make this assumption) that telling and teaching are the same thing: that what works in face-to-face conversation will work with a class of 20–30 students.

Unfortunately, we are not usually aware that, as a speaker in a face-to-face conversation, we depend on a system of nonverbal and verbal procedures that provide us with feedback about how our listener is understanding what we are saying. We use that feedback to try to head off misunderstandings and to constantly “repair” the conversation (Hutchby & Wooffitt, 1998). Such interactive repairing works effectively in one-to-one and small-group conversations when participants have a mutual interest in understanding each other but fails as the number of listeners or participants increases, their interest in understanding decreases, or the cultural distance between speaker and listener increases.

Although we know that teaching is more complex than conversing, we continue (for want of an alternative) to assume that the feedback we use in conversations (the eye contact, the facial expressions of interest, the nodding head, the questions) works equally well in the classroom. How else can we know whether students understand what we are saying?

Second, the practical aspects of teacher education work to sustain the common routines and rituals. Acquiring effective classroom practice is still largely based on an apprenticeship model. While theories and research are taught in university programs, the practical or applied aspects of teaching are learned “in the field” (Wideen, Mayer-Smith, & Moon, 1998). The practices and supporting beliefs that student teachers have been immersed in through their own experiences as school students are very resistant to change (Zeichner & Gore, 1990). When they enter a teacher education program, student teachers believe they already know how to teach and have high expectations about their ability to become good teachers (Weinstein, 1989). However, the “reality shock” of their first year of teaching throws them back into a preoccupation with classroom management (Livingston & Borko, 1989). Even when they plan to use new methods, they learn to cope by reverting to the methods by which they were taught (Segal, 1998). These involve the routines that they and their students accept and understand and that they experienced themselves as successful.

The reflective teacher model (Calderhead, 1989; Van Manen, 1977) that has come to play a significant role in thinking about teacher education and professional development for teachers also serves to sustain the ritualized routines and myths of teaching. When teachers reflect on their own practices, they can make use only of the evidence available to them. Since teachers are normally cut off from any direct evidence of student learning, they can only compare their awareness of their own classroom practices with the models they have in mind of “best practice.” Examples reported in

the literature suggest that these models are idealized forms of the standard routines supported by the commonly held myths that explain and sustain them (A. L. Cole, 1988; A. L. Cole & Knowles, 2000; Connelly & Clandinin, 1990).

External evaluation of teachers' competence is frequently based on observations of their classroom practices. When administrators evaluate a teacher's classroom performance, their judgments are primarily based on identifying the effective enactment of the ritualized routines of the busy, active classroom (Stodolsky, 1984). These are the routines that they understand, believe in, and expect to see. Teachers whose students play the reciprocal roles enthusiastically and with evident commitment are the teachers who are given awards for the quality of their teaching (Hopkins & Stern, 1996). Other teachers are invited to watch and copy what these award-winning teachers do.

Third (and probably more significantly) the ritualized routines 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 routines and the myths that support them. Empirical studies of the practices of the "best teachers" perform the same function. In most of 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 teachers. This kind of research provides some of the best, and most detailed, evidence we have of what these routines consist of and of the myths that sustain these routines (see, for example, Hopkins & Stern, 1996; Ramsay & Oliver, 1995). Yet the evidence from achievement test data shows that such reputations are inconsistent and frequently misplaced (Brophy & Evertson, 1973; Kirtman, 2002).

Perhaps the most significant evidence for the power of these myths to dominate both research methods and the research agenda is that I have never been able to find any competing evidence-based theories about how individual students' classroom experiences shape what and how they learn. There are many interpretations of what might be happening in students' minds (e.g., White & Frederiksen, 1998). But it does not seem a matter of interest to develop an evidence-based theory of how teaching is actually related to learning in mainstream research on teaching and learning.

A FINAL COMMENT ON THE ROLE OF RITUALIZED ROUTINES AND MYTHS

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 stay

within the ritualized routines of teaching. As teachers establish with their students, at the beginning of each year, their ways of using these routines, they become the basis on which teacher and students understand and can predict each other's expectations and actions. They give meaning to the many hours that students and teachers spend together, shaping and being shaped by each other's actions. The idealization of these routines and the wide acceptance of the supporting myths make it almost impossible for teachers to think or act otherwise. They have very little direct evidence, and little possibility of ever obtaining direct evidence, of what individual students in their classes are learning as they make those moment-by-moment decisions by which they run their classrooms (McNair, 1978).

Among other things, it makes it clear why attempts at reforming classroom teaching have so little long-term success (Elmore, Peterson, & McCarthy, 1996). Research has shown that after initial changes, teachers will "gravitate" back toward their former practices (Spillane, 2000) or that teachers will attempt to graft changes onto their current practices without changing the standard norms and routines (Coburn, 2002).

Ritualized social routines are a necessary part of effective social interaction, but the problem is that they become embedded in beliefs that are self-fulfilling and designed to sustain the routines independently of any evidence that they are achieving their intended goals. Commonly held beliefs about student ability are a very powerful way of explaining the effects of teaching by diverting attention away from the effectiveness of teaching routines and attributing differences in learning to characteristics of students. Standardized achievement testing is believed to produce valid information about what students know and can do. It serves to direct attention away from the ways students actually respond to testing situations.

It is important to search out independent evidence that the widely accepted routines of teaching are in fact serving the purposes for which they are enacted. We need to find a critical vantage point from outside the routines and their supporting myths. The approach taken by Stigler and Hiebert (1999; quoted at the beginning of this article) was to make comparisons across different cultures. The approach I have learned to take is to look at teaching through the eyes of students and to gather detailed data about the experiences of individual students. The account of my journey has been an account of the way this combination of objective and subjective research data has forced me to identify and question the myths by which we understand teaching and learning in schools.

But research is itself a system of rituals and beliefs. As many have pointed out, specific methods of research can shape perceptions and the kinds of questions that can be asked and answered (cf. Mosteller, Light, & Sachs, 1996). There are no absolute answers to this dilemma. In my own journey, I have tried to constantly question the purpose of my research and whether

the methods I was using served that purpose. The view I have reached is that the focus of research on teaching should be on direct observation of the realities of student experience and the processes that turn that experience into knowledge and skill. That does not just mean “giving students a voice” (A. L. Cole & Knowles, 2000). It means developing a precise, accurate, and replicable account of both the subjective and objective realities of student experience. In my view, the truth lies in the detail. Every generalization we make, every conclusion we draw must be true, in some important sense, of every individual. For example, there are many studies that aggregate data on student behavior, assuming that the student currently speaking or in the view of the observer represents all the students in the class (e.g., Herrenkohl, Palincsar, DeWater, & Kawasaki, 1999). Other studies aggregate student achievement data, assuming that individual differences in learning are of no significance (e.g., White & Frederiksen, 1998). If teaching practice is based on research that treats all students as an “average” student by aggregating student data or treats all teaching as a repetitive, changeless activity so that it can be validly represented by sample observations or recordings, then research will misrepresent teaching and student experience in ways that could be harmful to both teachers and students. The implication is that research needs to distinguish carefully between generalizations that are true of all individuals (such as my claim that basic conceptual learning processes are the same in both high- and low-ability students) and generalizations that are true only of some individuals in some contexts (e.g., my claim that spontaneous conversations about relevant curriculum content are more likely to occur among the more able students). To constantly work from the individual in specific contexts to the more general that is valid across contexts is what ethically responsible research on teaching must do.

As my journey indicates, it may take a very long time and a lot of data to begin to discover the reality beneath the cultural routines and myths by which we live and judge our performance (Nuthall, 1989).

Let me finish by describing a teacher who has attempted to step outside the standard routines of teaching and create a new kind of classroom. It illustrates how even innovative teachers can still be caught up in and misled by the dominant myths. Mrs. Middleton is an enthusiastic and energetic fifth-grade teacher. She believes students are better motivated if they can choose their own learning activities. Each year she prepares a wide range of resources and trains her students to plan and organize their use of these resources effectively. On one of the days that we observed and recorded in her classroom, the students were working on a science project on space. Each student had chosen her own research activity from a set of activities that Mrs. Middleton had carefully designed (using Bloom’s taxonomy) to challenge students’ research and thinking skills.

Rebecca was sitting at her table, surrounded by library books, looking for information about the differences between the surface features of Mars and the earth. Rebecca loves school. She especially loves this kind of project, and her research report was filled with accurate information and detailed drawings. Rebecca already knew a lot about space from the time when her father taught her (at 5 years old) to recite the names of all the planets in order.

Joel was sitting at his desk with one library book and the title of his project ("The Life of an Astronaut") written across the top of his report. He appeared to be busily searching the book for relevant information, but 20 min later he had written nothing underneath his title. Occasionally he went up to the teacher and complained about being interrupted by his neighbors (our observations showed these interruptions never occurred). On other occasions, he asked permission and disappeared from the room for 5 min or more. Joel spent long periods at his table just sitting and staring at his book. He never completed work. Whenever the teacher asked to see it, he had left it at home, accidentally lost it, or promised to complete it for homework.

Mrs. Middleton spends most of her time walking round the class monitoring progress and helping as required. Through the routines they have adapted to their classroom, Mrs. Middleton and her students know exactly what to expect of each other. Without such routines she could not manage 28 students simultaneously and know at a glance whether they were doing what they were supposed to be doing.

Mrs. Middleton believes Rebecca and Joel are serious students. Since neither asks for help, she satisfies herself that they are working well by glancing in their direction from time to time. She "knows" that Joel is a low-ability student and expects little of him. She is happy to see him reading a book. He knows what she expects of him and how to meet those expectations. Mrs. Middleton also "knows" that Rebecca is a high-ability student and expects to see her using several books simultaneously and writing a report that is many pages long. Rebecca understands these expectations and has learned how to meet them with the least effort.

Mrs. Middleton is a successful teacher. The students like her and try to please her. The principal is impressed with the efficient way she organizes and manages an individualized program. He proudly tells parents that she runs a program that caters to the individual needs of each student. But because Mrs. Middleton believes that the busy classroom is the learning classroom and that students' behavior is a function of their ability, she has no idea what Joel or Rebecca are actually learning. Joel is not reading any of the text in the book in front of him, but looking at the pictures. Rebecca knew all that she wrote in her report before this unit on space began. Day by day it is like this. The smoothly organized routines are played out. The students are constantly moving around, busy at their tasks. Mrs. Middleton

sees what she expects to see. Joel and Rebecca know how to meet her different expectations equally well. But Joel is headed for serious failure and Rebecca for outstanding success.

Mrs. Middleton was shocked when we showed her data from the achievement tests and interviews that revealed how the more able students had learned no more during the science unit than students she believed (on the basis of standardized test results) were less able. She assumed our evidence was somehow aberrant or mistaken. After all, she knew the students, she knew how hard they worked, she knew how successful her program was. Such is the power of culture and the routines and myths by which teachers structure and understand daily life in their classrooms.

I am grateful to the very large number of students, teachers, colleagues, graduate students and research assistants who have contributed to this journey. The credit is as much theirs as it is mine. The most recent research studies have been funded by the New Zealand Ministry of Education and the Marsden Fund.

Notes

1 The concept of culture has recently been strongly contested in anthropology for the role it has played in essentializing ethnic and other group differences. In this context I use the term in its traditional descriptive sense. An alternative might be the term *discourse* as it is used to represent a way of talking, thinking and acting.

2 More recent research has found inconsistent and conflicting relationships between teacher experience or expertise and student achievement (Borko, Livingston, McCaleb, & Mauro, 1998; Dewey, Husted, & Kenny, 2000; Verstegen & King, 1998). Part of the problem is that the relative effectiveness of teachers is unstable and varies with the kind of class they teach and the curriculum content they are teaching (Nuthall, 2004; Shavelson & Dempsey-Atwood, 1976; Stodolsky, 1984). Studies of the differences between expert and novice teachers are mostly based on teacher reputation which is known to be unrelated to student achievement (Sternberg & Horvath, 1995; but for an exception, see Borko & Livingston, 1989).

3 Please note that here (and later in this article) I have used the term *concept* as a shorthand for all the different kinds of knowledge, understandings, beliefs, and skills that are significant in areas like science, social studies, and mathematics.

4 Recently Richard Anderson and his associates (2001) have developed a procedure for tracking the cumulative changes that occur in students' use of argument stratagems in small-group discussions that they describe as the "snowball" effect.

5 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.

6 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).

7 These routines are similar to the patterned systems of behavior that anthropologists refer to as "rituals" (Quantz & Magolda, 1997). However, unlike rituals, many classroom routines are enacted unconsciously without the participants' being aware of the structured nature

of their behavior. It is like the way people use language without awareness of the complex grammatical structures that determine what they say and mean.

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, University of Canterbury, Christchurch, New Zealand.
- Alton-Lee, A. G., Densem, P. A., & Nuthall, G. A. (1991). Imperatives of classroom research: Understanding what children learn about gender and race. In J. Morss & T. Linzey (Eds.), *Growing up: The politics of human learning* (pp. 93–117). Auckland, New Zealand: Longman Paul.
- 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. Wellington: New Zealand Council for Educational Research.
- 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.
- Amerine, R., & Bilmes, J. (1988). Following instructions. *Human Studies*, 11, 327–339.
- Anderson, R. C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S. Y., Reznitskaya, A., Tillmanns, M., & Gilbert, L. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and instruction*, 19(1), 1–46.
- Ball, D. L., & Bass, H. (2000). Making believe: The collective construction of public mathematical knowledge in the elementary classroom. In D. C. Phillips (Ed.), *Constructivism in education: Opinions and second opinions on controversial issues* (99th Yearbook of the National Society for the Study of Education, Part I, pp. 193–224). Chicago: University of Chicago Press.
- Benenson, J., Apostoleris, N., & Parnass, J. (1998). The organization of children's same-sex peer relationships. In W. M. Bukowski & A. H. Cillessen (Eds.), *Sociometry then and now: Building on six decades of measuring children's experiences with the peer group* (New directions for child development, No. 80, pp. 5–23). San Francisco: Jossey-Bass.
- Bennett, N. (1976). *Teaching styles and pupil progress*. Cambridge, MA: Harvard University Press.
- Bohannon, P. (1995). *How culture works*. New York: Free Press.
- Bolster, A. S. (1983). Toward a more effective model of research on teaching. *Harvard Educational Review*, 53, 294–308.
- Borko, H., & Livingston, C. (1989). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American Educational Research Journal*, 26, 473–498.
- Borko, H., Livingston, C., McCaleb, J., & Mauro, L. (1988). Student teachers' planning and post-lesson reflections: Patterns and implications for teacher preparation. In J. Calderhead (Ed.), *Teachers' professional learning* (pp. 55–83). London: Falmer Press.
- Brophy, J. E., & Evertson, C. M. (1973). *Low-inference observational coding measures and teaching effectiveness*. Austin: University of Texas Research and Development Center for Teacher Education.
- Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York: Macmillan.
- Calderhead, J. (1989). Reflective teaching and teacher education. *Teaching and Teacher Education*, 5, 43–51.
- Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, 64(8), 723–733.
- Chan, D., Schmitt, N., Sacco, J. M., & DeShon, R. P. (1998). Understanding pretest and posttest reactions to cognitive ability and personality tests. *Journal of Applied Psychology*, 83(3), 471–485.

- 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.
- Coburn, C. E. (2002, August). *Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom*. Paper presented at the American Sociological Association Conference, Chicago.
- Cole, A. L. (1988). Personal signals in spontaneous teaching practice. *Qualitative Studies in Education*, 2(1), 25–39.
- Cole, A. L., & Knowles, J. G. (2000). *Researching teaching: Exploring teacher development through reflexive inquiry*. Needham Heights, MA: Allyn & Bacon.
- Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Harvard University Press.
- Connelly, F. M., & Clandinin, D. J. (1990). Stories of experience and narrative inquiry. *Educational Researcher*, 19, 2–14.
- Cuban, L. (1998). How schools change reforms: Redefining reform success and failure. *Teachers College Record*, 99(3), 453–477.
- Dahllöf, U. (1967). *School differentiation and teaching cycles*. Stockholm: Almqvist & Wiksell.
- 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.
- Derry, S. J. (1996). Cognitive schema theory in the constructivist debate. *Educational Psychologist*, 31(3–4), 163–174.
- 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.
- Doyle, W. (1986). Classroom organization and management. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 392–431). New York: Macmillan.
- Doyle, W., & Carter, K. (1984). Academic tasks in classrooms. *Curriculum Inquiry*, 14(2), 129–149.
- Dunkin, M. J., & Biddle, B. J. (1974). *The study of teaching*. New York: Holt, Rinehart, & Winston.
- Elbaz, F. (1983). *Teacher thinking: A study of practical knowledge*. New York: Nichols.
- Elmore, R. F., Peterson, P. L., & McCarthey, S. J. (1996). *Restructuring in the classroom: Teaching, learning, and school organization*. San Francisco: Jossey-Bass.
- Fischler, H. (1994). Concerning the difference between intention and action: Teachers' conceptions and actions in physics teaching. In I. Carlgren, G. Handal, & S. Vaage (Eds.), *Teachers' minds and actions: Research on teachers' thinking and practice* (pp. 165–180). London: Falmer Press.
- Flanders, N. A. (1970). *Analyzing teaching behaviour*. Menlo Park, CA: Addison-Wesley.
- Fodor, J. A. (1968). *Psychological explanation: An introduction to the philosophy of psychology*. New York: Random House.
- Gage, N. L. (1989). The paradigm wars and their aftermath: A "historical" sketch of research on teaching since 1989. *Educational Researcher*, 18(7), 4–10.
- Gagné, E., Yekovich, C. W., & Yekovich, F. R. (1993). *Cognitive psychology of school learning* (2nd ed). New York: Harper Collins.
- Gagne, F., & St. Pere, F. (2002). When IQ is controlled, does motivation still predict achievement? *Intelligence*, 30(1), 71–100.
- 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.
- Good, T. L., McCaslin, M. M., & Reys, B. J. (1992). Investigating work groups to promote problem solving in mathematics. In J. Brophy (Ed.), *Advances in research on teaching: Plan-*

- ning and managing learning tasks and activities, (Vol. 3, pp. 115–160). Greenwich, CT: JAI Press.
- Hallinan, M. T. (1996). Race effects on students' track mobility in high school. *Social Psychology of Education, 1*, 1–24.
- Harlow, A. (2000). *Student responses to international assessment items in science: A TIMMS example*. Unpublished master's thesis, University of Waikato, Hamilton, New Zealand.
- Herrenkohl, L. R., Palincsar, A. S., DeWater, L. S., & Kawasaki, K. (1999). Developing scientific communities in classrooms: A sociocognitive approach. *Journal of the Learning Sciences, 8*(3–4), 451–493.
- Hicks, D. (1993). Narrative discourse and classroom learning: An essay response to Egan's "Narrative and Learning: A Voyage of Implications". *Linguistics and Education, 5*, 127–148.
- Hoetker, J., & Ahlbrand, W. P. (1969). The persistence of recitation. *American Educational Research Journal, 6*(2), 145–167.
- Hogan, K., Nastasi, B. K., & Pressley, M. (2000). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition & Instruction, 17*(4), 379–432.
- Hopkins, D., & Stern, D. (1996). Quality teachers, quality schools: International perspectives and policy implications. *Teaching and Teacher Education, 12*(5), 501–517.
- 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.
- Hutchby, I., & Wooffitt, R. (1998). *Conversation analysis: Principles, practices and applications*. Malden, MA: Blackwell.
- Jackson, P. W. (1968). *Life in classrooms*. New York: Holt, Rinehart, & Winston.
- Kirtman, L. (2002). Policy and practice: Restructuring teachers' work. *Education Policy Analysis Archives, 10*(25). Retrieved from <http://epaa.asu.edu/epaa/v10n25/>
- Ladd, G. W. (1999). Peer relationships and social competence during early and middle childhood. *Annual Review of Psychology, 50*, 333–359.
- Lave, J. (1991). Situated learning in communities of practice. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 63–82). Washington, DC: American Psychological Association.
- Leinhardt, G., & Greeno, G. (1986). The cognitive skill of teaching. *Journal of Educational Research, 78*(2), 75–95.
- Livingston, C., & Borko, H. (1989). Expert-novice differences in teaching: A cognitive analysis and implications for teacher education. *Journal of Teacher Education, 40*(4), 36–42.
- McLaren, P. (1993). *Schooling as ritual performance: Towards a political economy of educational symbols and gestures* (2nd ed.). London: Routledge & Kegan Paul.
- 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, UK: Penguin.
- Mosteller, F., Light, R., & Sachs, J. (1996). Sustained inquiry in education: Lessons from skill grouping and class size. *Harvard Educational Review, 66*, 797–828.
- Nunan, D., & Lamb, C. (1996). *The self-directed teacher: Managing the learning process*. New York: Cambridge University Press.
- 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). Chicago: Rand McNally.
- 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. (1978, March). *Classroom processes in the Second IEA Mathematics Survey*. Paper presented at the annual meeting of the American Educational Research Association, Toronto.

- Nuthall, G. A. (1989). How can research on teaching help teachers. In *NZARE 1989 Conference Proceedings: Keynote Addresses* (pp. 33–60). Wellington: New Zealand Association for Research in Education.
- 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, The Netherlands: Kluwer Academic.
- 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), 141–256.
- 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. (2000c). The role of memory in the acquisition and retention of knowledge in science and social studies units. *Cognition & Instruction*, 18(1), 83–139.
- Nuthall, G. A. (2001a). Understanding how classroom experience shapes students' minds. *Unterrichts Wissenschaft: Zeitschrift für Lernforschung*, 29(3), 224–267.
- Nuthall, G. A. (2001b, 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. (2001c, 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 Ninth Biennial Conference of the European Association for Research on Learning and Instruction, Fribourg, Switzerland.
- Nuthall, G. A. (2004). Relating classroom teaching to student learning: A critical analysis of why research has failed to bridge the theory-practice gap. *Harvard Education Review*, (Fall), 274–306.
- 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. (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. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.
- 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.
- Peterson, P. L., & Walberg, H. J. (Eds.). (1979). *Research on teaching: Concepts, findings, and implications*. Berkeley, CA: McCutchan.
- Quantz, R. A., & Magolda, P. M. (1997). Nonrational classroom performance: Ritual as an aspect of action. *Urban Review*, 29(4), 221–238.

- Ramsay, P., & Oliver, D. (1995). Capacities and behavior of quality classroom teachers. *School Effectiveness and School Improvement*, 6, 332–366.
- Rogoff, B. (1993). Children's guided participation and participatory appropriation of socio-cultural activity. In R. H. Wozniak & K. W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments* (pp. 121–153). Hillsdale, NJ: Erlbaum.
- Roth, W.-M. (2001). Situating cognition. *Journal of the Learning Sciences*, 10(1–2), 27–61.
- Sahlstrom, F., & Lindblad, S. (1998). Subtexts in the science classroom: An exploration of the social construction of science lessons and school careers. *Learning and Instruction*, 8(3), 195–214.
- Sammons, P., Hillman, J., & Mortimer, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research*. London: International School Effectiveness and Improvement Centre, Institute of Education, University of London.
- Sanchez, R. J., Truxillo, D. M., & Bauer, T. N. (2000). Development and examination of an expectancy-based measure of test-taking motivation. *Journal of Applied Psychology*, 85(5), 739–750.
- Segal, S. (1998). The role of contingency and tension in the relationship between theory and practice in the classroom. *Journal of Curriculum Studies*, 30(2), 199–206.
- Shanks, J. (1994). Student reactions to a standardized curriculum. *Journal of Curriculum and Supervision*, 10, 43–60.
- Shavelson, R., & Dempsey-Atwood, N. (1976). Generalizability of measures of teaching behavior. *Review of Educational Research*, 46, 553–611.
- Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means of studying cognitive development. *American Psychologist*, 46, 606–620.
- Simon, A., & Boyer, G. E. (Eds.). (1968). *Mirrors for behavior: An anthology of classroom observations instruments*. Philadelphia: Research for Better Schools Inc.
- Skowronek, H., & Schmied, D. (Eds.). (1977). *Forschungstypen und Forschungsstrategien in der Erziehungswissenschaft*. Hamburg, Germany: Hoffmann & Campe.
- Spillane, J. P. (2000). Cognition and policy implementation: District policymakers and the reform of mathematics education. *Cognition and Instruction*, 18(2), 141–179.
- Spindler, G. (1999). Three categories of cultural knowledge useful in doing cultural therapy. *Anthropology and Education Quarterly*, 30(4), 466–499.
- Sternberg, R., & Horvath, J. (1995). A prototype view of expert teaching. *Educational Researcher*, 24, 9–17.
- 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* (NCES 99-074). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.
- Stodolsky, S. (1984). Teacher evaluation: The limits of looking. *Educational Researcher*, 13(9), 11–18.
- 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.
- Van Manen, J. (1977). Linking ways of knowing with ways of being practical. *Curriculum Inquiry*, 6, 205–208.
- 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.

- Wainer, H., & Braun, H. I. (Eds.). (1988). *Test validity*. Hillsdale, NJ: Erlbaum.
- Wang, J. (1998). A content examination of the TIMMS results. *Phi Delta Kappan*, 80(1), 36–38.
- Weinstein, C. S. (1989). Teacher education students' preconceptions of teaching. *Journal of Teacher Education*, 40(2), 53–60.
- Wells, G. (2002). *Learning and teaching for understanding: The key role of collaborative knowledge building*. In J. Brophy (Ed.), *Social constructivist teaching: Affordances and constraints* (Advances in research on teaching, Vol. 9, pp. 1–41). Oxford, UK: Elsevier Science.
- Wentzel, K. R. (1999). Social-motivational processes and interpersonal relationships: Implications for understanding motivation in school. *Journal of Educational Psychology*, 91, 76–97.
- White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modelling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, 16, 3–118.
- Wideen, M., Mayer-Smith, J., & Moon, B. (1998). A critical analysis of research on learning to teach: Making the case for an ecological perspective on inquiry. *Review of Educational Research*, 68(2), 130–178.
- 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.
- Zeichner, K. M., & Gore, J. M. (1990). Teacher socialization. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 329–348). New York: Macmillan.

GRAHAM NUTHALL was Professor Emeritus in Education at the University of Canterbury. He completed a doctorate at the University of Illinois and was involved in research on teaching and learning in classrooms in the United States and New Zealand for more than 45 years. Recent reports of this research can be found in “Learning How to Learn: The Evolution of Students' Minds through the Social Processes and Culture of the Classroom” (*International Journal of Educational Research*, 1999) and “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*, 2000). Professor Nuthall passed away in July 2004.