Group Work for the Good
Unpacking the Research behind One Popular Classroom Strategy

By Tom Bennett

It wasn’t until I had been teaching 11- to 18-year-olds for four years that I realized I had been consistently misled. Up until that point I had trusted my teacher training to provide the best of what had been discovered in the discipline of teaching and learning. If I had been shown a method or theory by which I could perform my job more efficiently, I assumed it would have been forged in the crucible of experience and evidence. I assumed that what we knew about teaching, say, chemistry, for example, progressed in a linear, accumulative way. But I found the opposite.

As a philosophy and religious studies high school teacher in the United Kingdom, I discovered that a good deal of what was considered orthodoxy in my profession was unsubstantiated. I believe many of my teacher colleagues in the United States have made similar discoveries.

In 2004, I had just emerged from the U.K. Department for Education’s Fast Track recruitment program into teaching, where I had spent weekends learning about Neuro-Linguistic Programming, a program called Brain Gym, and how to sort my students according to their learning styles.* I was told that my students possessed multiple intelligences, and it was strongly hinted to me that the more technology I could accommodate into my lessons, the better their needs as digital natives would be met. My initial classroom design of rows and columns was frowned upon, and tables and horseshoes were recommended. And all because, I was told, the research confirmed each avenue.

Except that it didn’t. Often, it barely addressed the topics. I won a teacher fellowship at Cambridge University, where I was given the opportunity to pull back the curtain of the mighty Oz of research. It was an epiphany. As I learned to navigate the university’s endless libraries of education journals and papers, I was struck by a thought that at first I dismissed as impertinence: a good deal of research I had been recommended as a new teacher

was astonishingly misguided.‡ I felt like a heretic. The temerity of my emergent conclusion struck me as astonishing, right? But my master’s degree in philosophy (with a focus on epistemology) kept pointing me back the same way: a lot of what was considered research was often based on little more than bias or opinion.

For instance, Neuro-Linguistic Programming was a ragbag of fashionable pseudoscience that had been broadly criticized, even at the time of its publication, but still it lurched on for decades. It was mystical hoo-ha that rested on the “science of success” that predicted among other things that you could tell when someone was lying. Learning styles had similarly been dead on a mortuary slab for many years, but even today teachers are earnestly instructed in their use. Brain Gym was, until recently, considered to be cutting-edge practice, including the claim, widely believed by Brain Gym enthusiasts, that water should be held to the roof of the mouth because it reaches the brain quicker that way. And so on. Everywhere you looked, education was, and is, deviled by what physicist Richard Feynman would call cargo cult science, aping the form of science in every way but the ones that mattered.

It inspired me to write Teacher Proof: Why Research in Education Doesn’t Always Mean What It Claims, and What You Can Do about It, on which this article is based. My book bluntly exposes some of the bigger education myths that still rattle their chains in the classroom. Each chapter is devoted to a questionable educational theory; I examine each claim made for its efficacy by simply following the research crumbs backward.

Often, I found that a claim would refer to this or that seminal paper, which I would then find rested its evidence base on some other seminal paper, which I would then pursue and so on. Madly enough, I often found the most basic of problems: papers that referred to works by the same author, papers that relied on the most minuscule of sample sizes, papers that failed to in any way test their own hypothesis to failure, and so on. I found enormous over-reliance on opinion and testimonials as proof of any kind.

In short, I found what you might find in any science, but it seemed to be magnified in educational science. Why? One reason was that social science practitioners frequently proposed that what might be classed as proof in their field did not have to meet the rigors of the physical sciences, which is understandable given the challenge of dealing with human beings, who are not inert objects of examination but rather can be difficult and interactive participants in their own analysis. But instead of acknowledging this profound obstacle, many researchers simply ignore it.

There is a great deal of excellent research in education, but it is often drowned out by the cacophony of the fashionable, the exciting.

In 2012, to help remedy this disastrous state of affairs, I founded researchED (www.workingoutwhatworks.com), a teacher-led, grass-roots wiki movement aimed at empowering teachers through greater research literacy and bringing together the best research for the classrooms that need it most. Since our first conference, it’s taken off, and now we’re preparing for a researchED conference in New York in May 2015. Clearly, there’s an appetite among many teachers to no longer be beholden to the institutions responsible for their support, and instead to find out—through a process of profound reprofessionalization—what actually works.

One of the most enduring myths I’ve encountered in education is the subject of this article: group work. I’ve seen entire educational districts seized by the belief that group work is the only way for students to learn, or at least by far the most efficient way. I spent years wrestling with the tension between the claims supporting this teaching method and the evidence of my own classrooms. And when I investigated the foundations on which these claims were made, I found that they were often not substantiated in any credible way.

Group work does have its place in the classroom. Allowing students to partner on a particular assignment can engage them in the subject matter they are studying, help them improve their skills, and teach them the value of teamwork—as long as the students, themselves, do one crucial thing: stay on task.

Group work can go a long way in reinforcing content knowledge. But it should not take the place of fully guided instruction.

When students bring the necessary focus to group work, and when teachers use it appropriately—that is, to supplement instruction, not replace it—group work can go a long way in reinforcing content knowledge. But it should not take the place of fully guided instruction, which sound research (not the kind I discuss above) has overwhelmingly found is most effective in helping students learn.§ Still, in recent years, group work has become one of many fads to seemingly conquer the education world. Why is this so?

Where Did Collaborative Learning Come From?

In the early 2000s, a growing swell of research seemed to support the use of group work as one of the best ways to learn.1 Proponents claimed the strategy would:

- Improve learning;
- Develop social skills;
- Develop empathy and altruism;
- Deepen learning;
- Improve test scores and retention;
- Develop complex learning strategies;
- Create independent learning; and
- Enable lifelong learning.

---

1For more about how to tell good research from bad in education, see “Measured Approach or Magical Elixir?” in the Fall 2012 issue of American Educator, available at www.aft.org/ae/fall2012/willingham.

§For more about fully guided instruction, see “Putting Students on the Path to Learning” and “Principles of Instruction” in the Spring 2012 issue of American Educator, available at www.aft.org/ae/spring2012.
It was also widely reported at the time that group work, or collaborative learning, was widely misused; students often sat together but infrequently learned together in a meaningful way. There were many other reports at this time, and many more afterward, all saying the same thing: if you want children to learn, then they need to be in pods. In other words, the research seemed to say that group work worked, and if it didn’t, you weren’t doing it right.

Very early on in my career, this was one of the pieces of absolutely infallible dogma I had been told to adopt as a way of driving learning. In one of my first few years as a teacher, I was observed teaching a lesson which, while surely not perfect, was rated unsatisfactory. When I queried it, I was told that because there was no group work, the students couldn’t really be learning deeply enough. Regardless of outcome for the students, the process had predetermined the evaluation, almost as if the winner of the 100-meter race at the Olympics had been decided by the athlete who most closely conformed to the preferred sprinting style.

When you’re a rookie and you don’t have the confidence to question authority, that kind of judgment is a punch to the gut. It ruined me for months, as professional criticism often can. My lesson wasn’t just judged to be average and bland, it was below par—it had failed. I was failing my students.

We can see the modern incarnation of group work emerging from such theorists as Lev Vygotsky and Jean Piaget. Vygotsky, an early 20th-century Russian psychologist, has been a major influence in the past few decades. He believed that social interaction precedes development; action is the basis of forming thoughts. According to his child-centered understanding of how we learn, pupils occupy the roles of problem solvers, and teachers are there as facilitators; this is the famous transition from the sage on the stage to the guide from the side. Language used by children is a tool used in order to think. Talk, for Vygotsky, is a learning tool. He believed that the use of talk—group work, discussion—in the classroom would help to reduce the pupil’s “zone of proximal development” or the gap between where he or she could be and his or her current stage of learning. As Vygotsky put it: “What a child can do today in cooperation, tomorrow he will be able to do on his own.... The students are responsible for one another’s learning as well as their own. Thus, the success of one student helps other students to be successful.”

Proponents of collaborative learning claim that the active exchange of ideas within small groups not only increases interest among the participants but also promotes critical thinking. As Anuradha Gokhale stated in a 1995 article, “There is persuasive evidence that cooperative teams achieve at higher levels of thought and retain information longer than students who work quietly as individuals. The shared learning gives students an opportunity to engage in discussion, take responsibility for their own learning, and thus become critical thinkers.”

According to Vygotsky, students are capable of performing at higher intellectual levels when asked to work in collaborative situations than when asked to work individually. Group diversity in terms of knowledge and experience contributes positively to the learning process. Psychologist Jerome Bruner contends that cooperative learning methods improve problem-solving strategies because the students are confronted with different interpretations of the given situation. The peer support system makes it possible for the learner to internalize both external knowledge and critical-thinking skills and to convert them into tools for intellectual functioning.

The theories that underpin collaborative learning can broadly be described as deriving from constructivism, the idea that we are active participants in the process of learning, not passive recipients of experience and factual accumulation. We construct new meaning as we encounter experiences and build our learning on what we already know. This means that learning is intrinsically a social process and is inseparable from the process of interacting with others. Constructivist thinking can be found in the works of such educational pioneers as Maria Montessori, Piaget, John Dewey, and Bruner. Constructivism is not by itself an educational method but a description of how constructivists believe learning takes place, whatever method is used. But advocates of constructivism have derived pedagogies from its principles, such as active learning, discovery learning, and collaborative learning, the aim being, of course, to exploit the natural process of learning by artificially reproducing the best environments in which it takes place.

The problem that immediately arises is that this conceptual model of how learning occurs is contested. One challenge it faces is that it is unverifiable. How would we know if such a theory was true or false? What evidence would invalidate it? We see this problem reoccur frequently throughout literature that supports the collaborative learning model: there are many detailed blueprints about what it means to learn collaboratively but less indication that these blueprints correspond to a meaningful description of the actual learning process.

One of the advantages of group work, the research assures us, is that it’s a great way to engage children. Well, that seems to be true, if by engage you mean give them a chance to do less for a period of time and catch up with each other. Of course, many kids will leap like salmon into that river; kids really do like working in groups. But the point of group work is that it is supposed to develop and encourage skills of interactivity and motivation. In the examples where it seems to work best, those qualities and skills have to pre-exist the activity.

I must emphasize that I’m not against group work, and that I use it myself when I want students to practice knowledge recall with each other, or when I want to change the pace of a lesson full of direct instruction. I enjoy it, especially with upper high school students, who can produce some astonishingly good work (for exam-
ple, through fantastic debates and discussions that unpack prejudice and challenge axioms, or resources that the whole class can use later on like posters and notes). I use anything that gets results, and I’ll try anything that looks like it works for other people. But the insistence that group work is the best way to develop higher-level thinking skills, and that it has an appreciable, improving effect on students overall, is just undemonstrated. One paper I read that celebrated its benefits had 48 participants. That’s forty-eight. I could fit them in my tiny British garden. I’d barely call that evidence. It’s a pattern we see time and time again in poor education research: tiny samples, short study intervals, and muscular, hopelessly optimistic extrapolation from a microscopic set of data points. One such study barely qualifies as research. Several studies, all reproducing the same flaw, still don’t constitute an evidence base, for reasons that are obvious. It’s also a problem with meta-studies in this area.

What Does the Research Say?
“There is an ever increasing need for interdependence in all levels of our society today. Providing students with the tools to effectively work in a collaborative and cooperative environment should be our priority as teachers. Cooperative learning (CL) is one way to provide students with a well defined framework from which they can learn from one another.”

That was from the very opening paragraph of an online paper to which I was referred by a group-work enthusiast. First sentence: unproven and unprovable conjecture, opinion, and subjective values of the author. It doesn’t really bode well for the rest of the paper.

Another paper I looked at included 250 students—not a large sample by any means. The study was focused on video-recorded evidence of group work after months of group-work training for both teachers and students. Of course, permission had to be obtained for filming, and when it wasn’t, the pupils were removed from the test subjects. Students knew that they might be filmed that day. Teachers probably did. Groups of students were then given group tasks designed to display problem-solving ability, cooperativeness, etc. So the tasks themselves were factors in the process; what might the researchers have found in tasks that weren’t designed to show the quality tested?

Researchers then had to watch selections of the clips, and decide to what degree students were on task and engaged, and what kind of quality of engagement they displayed. These are tremendously subjective properties and could vary from researcher to researcher, from day to day, depending on a million factors, subconscious and not.

And what did they find? The test groups displayed better group-work skills than the ones who had not been through the training process. They found exactly what they wanted to find, and given the way they loaded the dice from the outset, I’m not surprised. Because they used something very common in successful evaluations of contested objects of research such as collaborative learning: loaded proxy indicators.

It’s easy enough to measure height or temperature. We have tape measures and thermometers for such things. But how do we measure something more abstract, like learning? What we do is try to capture the next best thing: something that we can measure...

The (Usually Ignored) Drawbacks of Group Work
The claims made by advocates of group work are frequently utopian, because in a real classroom many variables work against the success of any group activity. These include:

1. **Disguised inactivity.** In group scenarios, students are provided with an opportunity to really put their backs into doing nothing. If you give a task to three or four people, one or two may realize it’s time to freeze, because others will carry the burden of the task. In the meantime, they can coast under the guise of “research” or “running the group.” Their inaction is hidden inside the smog of collaborative effort.

2. **Unequal loading.** Related to this is the problem that while every student might participate, the participation might be profoundly uneven. Some will contribute at glacial speeds, while others will race and caper through every task and subtask.

3. **Inappropriate socialization.** Students may end up competing to see who can discuss the task the least. Playtime has come early in this scenario. Pupils are well aware that group work can devolve into recess.

4. **Unfair assessment.** When I praise a pupil, it’s a clear one-to-one relationship. In grading groups, we often must give collective grades. We should do this as rarely as possible, and praise and reward individual effort where possible. Groups, after all, cannot think or learn; that is possible only for individuals.

One of the main tasks of the teacher is to introduce children to the best of what has already been discovered and thought.
that we think will correspond with the quality we’re interested in. For example, we can’t see electrons, but we know that whenever you get them, you have voltage and amps and electricity, and light bulbs glow.

In this experiment, what did they use as proxy measures? I’ve mentioned that they noted degree of engagement, quality of interactivity, sustained levels of discussion, and numerous other factors. These things could at least be observed by the researchers with their senses. They found that at the end of the experiment, test groups of pupils had longer discussions, maintained their groups better, and interrupted each other far less than in the control groups.

This somehow proves that group work improves learning? Or maybe it proves that groups trained at group work get better at group work. Or it could prove a million other things. Or nothing.

I find group work useful when pupils have spent long periods in private work, individually and self-managed.

That’s the problem. We don’t know. And neither do the researchers, who designed an experiment with a success criteria that revolved around “being better at working in groups.” Working in groups helps you improve working in groups, apparently.

One of the recurrent themes in the literature about group work was the claim that students and classrooms would benefit from group work if they were trained in the skills necessary to interact as groups. This is the cart before the horse; if effort can be put into ensuring that children can behave well enough to participate meaningfully in group work, then that same effort can be directed toward teaching them just as well in non-group environments.

In 2006, both the BBC and the Guardian reported research that claimed that schools were failing to implement group work effectively. This research, backed up by a much larger study of 4,000 students over a year in grades 1–9, seemed to testify to the same claims made elsewhere: pupils working in groups collaborate more, learn more, socialize more, and are more motivated to succeed. But motivation was measured by the proxy of self-evaluation questionnaires, which is a notoriously bad way of ascertaining the truth, as you will find when you survey people about their history of honesty, drug taking, and other patterns of behavior. Self-reported surveys aren’t meaningless, but they’re a long way from the level of data we need to evaluate the efficacy of any one learning method.

The paper referred to above generously mentions that the Hawthorne effect (a well-documented phenomenon where participants in an experiment subconsciously attempt to meet the intended outcomes of the study) may be a factor, but then fails to explore it beyond saying that the researchers tried their best to keep it to a minimum, and that they believed it wouldn’t have much of an effect. And they conclude that they have proven students who go through the program probably get better at the kind of things the program is designed to teach. The website accompanying the paper claims that “Experimental research on small groups and psychological theory emphasises that effective group work in classrooms has enormous potential in terms of increasing children’s motivation and learning.” The researchers appear to start from a hypothetical premise that doesn’t even seek to explain an observed phenomena, but rather to confirm it, with an aim to rolling it out across mainstream schools.

The Opportunity Cost of Group Work
Group work also presents us with another problem: opportunity cost. What are the students not doing when they are doing group work? Is it the most effective and efficient use of time, which is one of the most precious resources they have access to? Or could they be doing something else, with more impact?

Take, for example, when pupils are split into groups and given “seek and return” missions with specific learning goals. That’s fine, but it is incredibly time consuming; half a lesson can easily be consumed in the conveyance of a group of facts that could be far more efficiently conveyed in five minutes.

Group work, to put it simply, takes a long time, and the knowledge it conveys could often be far more efficiently imparted through other methods, such as direct instruction. This isn’t to say that learning is simply the accumulation of facts, but rather, that the process of that accumulation is far less well facilitated through group work.

Another conceptual problem: the idea that children learn best from other children; that they are the sources of all the information they need. This isn’t a bad idea when it comes to getting them to think about alternatives and ideas and values opposite their own, because one student’s opinion about something is just as good as another’s for learning about justifications and difference. But when it comes to factual conveyance, that’s what a subject expert is for. For every subject, there is an enormous body of content that is beyond dispute, even within the humanities, and that is one of the main tasks of the teacher—to introduce children to the best of what has already been discovered and thought. If we don’t do that, we break the link between children and the legacy of our ancestors. You might as well start from scratch. That’s not something I want to do with my students. I want them to build on what I and others have learned, and hopefully, to surpass us. I refuse to hobble them by forcing them to discover everything for themselves all over again.

Finally, there’s a problem associated with classroom management, often completely overlooked by those who advocate group work: it can wreak havoc on the behavior in your classroom. The temptation for children to be off task is simply too great for many of them, and I witness new teacher after new teacher struggle before a class of kids who are all facing each other and not the front of the class or the teacher. It’s an invitation to misbehave. Many of the studies I have read have been conducted in schools that could be best described as pleasant, with groups of students who are best
described as amenable.

But group work in a difficult class, with a new teacher, can be chaos. You won’t see much independent effort and collaborative thinking then. Or indeed, learning. I coach a lot of new teachers who are struggling with their students’ behavior. And one of my first pieces of advice: hold off on the group work until you can manage your classroom.

**When Is Group Work Useful?**

My intent isn’t to discredit group work just because some have made exaggerated claims; it’s a perfectly sound approach in the classroom for many activities. Sadly, to find good research that offers a more cautious approach to the best use of group work, it is often necessary to step outside of the field of pure educational research, and into cognitive psychology or business, where more sober research has been done. Former Harvard psychology professor J. Richard Hackman, for example, is worth reading on the broader issues behind teams. In schools, there are several good reasons to do group work:

- **p** In situations where tasks are impossible to achieve without it—for example, football or an orchestra.
- **p** To vary the type of classroom activity—for example, moving from a period of individual book work to a short session of cooperation, in order to stimulate the pupils by the ancient method of mixing things up a bit. A change is as good as a rest and so on.
- **p** To improve students’ ability to cooperate, reason with each other, listen to others’ opinions, and so on.

These are some of the more common reasons espoused by advocates of group work. They are valuable goals for child development, and I’m happy to use group work from time to time as one way to support that.

It is possibly almost too obvious to say, but every teacher should have an aim for what he or she wants to achieve in his or her lesson, even if that aim changes as the lesson progresses. The prudent teacher then attempts to match the student activity with the method that strikes him or her as being most appropriate to achieving that aim. Utility should be the heart of this decision-making process.

I find group work useful when pupils have spent long periods in private work, individually focused and self-managed. This form of directed learning is fine for stretches, but the human mind palls at repetition and monotony. If, as Aristotle claims, “man is a social animal,” then the wise teacher has mercy on the attention spans of young minds and allows cramped muscles to stretch. There is value in discussion between pupils when it can be guaranteed that task focus can be maintained. Ability grouping can produce a variety of interesting outcomes: more able students can push each other to new heights, and mixed-ability groupings can allow the more able to coach the less able and provide the teacher with an army of assistants.

In my own classroom, when I am sure that my older pupils have understood the content of a philosophy unit, I consolidate the learning through group work. Let’s say we have worked through a unit on Kantian deontology. I then pair off the students (in groups of my choosing) to argue, back to back, for and against a simple motion regarding Kant’s ethical theories. After two minutes, I blow a whistle and they have to reverse their positions conceptually, arguing against their previous positions. Then I blow again, and they reverse. It’s a powerful activity that achieves an end that could not be easily achieved individually. It is hard to sharpen a knife against itself. But against a stone, or another knife?

After that, I get them to build a poster with as many arguments for and against Kant’s position as they can think of. They then pass these posters to the next group of students, who correct or steal any points they can. The posters carry on until they return home to their original groups. Finally, I give the students 20 minutes to prepare for a formal debate, with a motion and groups created by me, with rules of conduct and scoring. All of these activities are ideal in collaborative forms; they use students to drive each other’s recall and force them to make connections between points of understanding. Note that this requires the students to understand the content prior to the execution of the tasks; the group activities support fluidity and consolidation, not excavation.

Here’s my parting advice: use group work when you feel it is appropriate to the task you want your students to achieve, and at no other time. The irony of the advocates’ position is that while it correctly identifies the many benefits to using group work, their error is made when group work is preferred over other strategies because of some imagined potency, or when it is fetishized as a method imbued with miraculous properties. It isn’t dogma, it isn’t a panacea, and it isn’t the messiah. It’s one strategy among many. And it’s a perfectly reasonable part of a teacher’s arsenal of strategies. Not because pseudo-research has settled the matter, but because the teacher feels it appropriate at that time, for that lesson, with those children. And not before.

(Endnotes on page 44)
Group Work
(Continued from page 37)

Endnotes


