EDUCATION: FUTURE FRONTIERS

OCCASIONAL PAPER SERIES

How to Teach Critical Thinking

Daniel T. Willingham

A paper commissioned by the NSW Department of Education

ABOUT THE AUTHOR

Daniel T. Willingham earned his B.A. from Duke University in 1983 and his Ph.D. in Cognitive Psychology from Harvard University in 1990. He is currently Professor of Psychology at the University of Virginia, where he has taught since 1992. Until about 2000, his research focused solely on the brain basis of learning and memory. Today, all of his research concerns the application of cognitive psychology to K-16 education.

He writes the "Ask the Cognitive Scientist" column for American Educator magazine, and is the author of *Why Don't Students Like School?*, *When Can You Trust the Experts?*, *Raising Kids Who Read*, and *The Reading Mind*. His writing on education has appeared in sixteen languages.

In 2017 he was appointed by President Obama to serve as a Member of the National Board for Education Sciences.

EDUCATION: FUTURE FRONTIERS

© Daniel T. Willingham and the State of New South Wales (Department of Education), 2019.

EDUCATION: FUTURE FRONTIERS is an initiative of the NSW Department of Education exploring the implications of developments in AI and automation for education. As part of the Education: Future Frontiers Occasional Paper series, the Department has commissioned essays by distinguished authors to stimulate debate and discussion about AI, education and 21st century skill needs. The views expressed in these essays are solely those of the authors.

The Australian Curriculum acknowledges that developing thinking skills is a primary purpose of education, and identifies critical thinking as an important capability for the 21st century. Critical thinking has, of course, long been a valuable skill for young people to master, though its importance is expected to increase as the world becomes ever more augmented by artificial intelligence and other emerging technologies. Despite consensus on the need for critical thinking, there is still considerable debate over how it is learned and, subsequently, how education can best support students to develop critical thinking capabilities. Some believe that critical thinking can be taught as a generic skill independently from subject content, while others contend that content mastery is pivotal to the development of thinking capabilities. This paper considers what cognitive science can tell us about how critical thinking is acquired, and the implications for how education might best develop young people's critical thinking capabilities in light of this evidence.

The author concludes that scientists are united in their belief that content knowledge is crucial to effective critical thinking. Scientists are somewhat divided as to whether critical thinking is best characterised as a large number of more specific skills or a smaller number of more generic skills. The author argues that the former is not a fruitful way to conceptualise skills in education, however, as there is little theory to guide how to teach generic skills. The author recommends a four-step process to develop a program to teach critical thinking: (1) identify a list of critical thinking skills for each subject domain; (2) identify subject matter content for each domain; (3) plan the sequence in which knowledge and skills should be taught; (4) plan which knowledge and skills should be revisited across years.

Individuals vary in their views of what students should be taught. How should teachers discuss misdeeds of a nation's founders? What is the minimum accomplishment expected of each student in mathematics? But there is no disagreement on the importance of critical thinking skills. In free societies, the ability to think critically is viewed as a cornerstone of individual civic engagement and economic success. We may disagree about which content students should learn, but we at least agree that, whatever they end up learning, students ought to think critically about it.

Despite this consensus it's not clear we know what we mean by "critical thinking." I will offer a commonsensical view (Willingham, 2007). You are thinking critically if (1) your thinking is novel—that is, you aren't simply drawing a conclusion from a memory of a previous situation and (2) your thinking is self-directed—that is, you are not merely executing instructions given by someone else and (3) your thinking is effective—that is, you respect certain conventions that make thinking more likely to yield useful conclusions. These would be conventions like "consider both sides of an issue," and "offer evidence for claims made," and "don't let emotion interfere with reason." This last characteristic will be our main concern, and as we'll see, what constitutes effective thinking varies from domain to domain.

An alternative informal definition holds a different characteristic of thinking as key: thinking when others might not. For example, if you want a long black at your local cafe, you would probably just order it and pay your three dollars. But you might notice that the shop charges 35 cents for hot water and 75 cents for an espresso shot added to any drink; you could order hot water and a shot instead. What makes this example interesting is that someone could think to try working the angles of a coffee shop menu whereas most people would not. It's not the difficulty of thinking successfully, it's deciding to think in the first place. Educators hope to instil this quality

in students; we want them to question articles they read in the media for example, or to think through whether the claims of an advertisement make sense. This appetite for cognitive work when others might avoid it seems to be partly a matter of personality (Cacioppo et al., 1996). It may be educable, but there's limited research on the matter.

This paper will focus, then, on the first sense in which educators use the term critical thinking, namely, successful thinking. Of course we want students to choose to think, but we won't be satisfied if their thinking is illogical, scattered, and ultimately fails. Teaching critical thinking that succeeds has been the subject of considerable research. The remainder of this paper reviews important insights of this research, and closes with recommendations as to how these findings can inform the teaching of critical thinking.

CRITICAL THINKING CAN BE TAUGHT

Planning how to teach students to think critically should perhaps be our second task. Our first should be reassuring ourselves that such instruction is needed and can succeed. Perhaps learning to think critically is akin to learning language as an infant. In a language-rich environment and with frequent situations where it is useful, the child will learn to use language without any formal instruction. Perhaps in the same way, you learn about critical thinking based on what's available to you in the environment. Is there evidence that explicitly teaching critical thinking brings any benefit?

There is, and such evidence is available for different subject matters. For example, in one experiment researchers taught college students principles for evaluating evidence in psychology studies—principles like the difference between correlational research and true experiments, and the difference between anecdote and formal research (Bensley & Spero, 2014). These principles were incorporated into regular instruction in a psychology class, and their application was practiced in that context. Compared to a control group that learned principles of memory, students who learned the critical thinking principles performed better on a test that required evaluation of psychology evidence.

There is even evidence that critical thinking skills can be taught and applied in complex situations under time pressure. In one experiment, officers in the Royal Netherlands Navy received training in the analysis of complex battlefield problems in a high-fidelity tactical simulator. They were first taught a sequence of steps to undertake when analyzing this sort of problem, and then underwent a total of 8 hours of training on surface warfare problems, with feedback from an expert. The critical outcome measure was performance (without feedback) in a new surface warfare problem, as well as performance on air warfare problems. Judges assessed the quality of participant's action contingency plans, and those receiving the training outperformed control subjects (Helsdingen et al., 2010).

There are many other examples of critical thinking skills that are open to instruction (Abrami et al., 2008; Bangert-Drowns & Bankert, 1990). But perhaps we should not find this result terribly surprising. You tell students that this is a good strategy for this type of problem, and you have them practice that strategy, so later they use that strategy when they encounter the problem.

PLANNING HOW TO TEACH STUDENTS TO THINK CRITICALLY SHOULD PERHAPS BE OUR SECOND TASK. OUR FIRST SHOULD BE REASSURING OURSELVES THAT SUCH INSTRUCTION IS NEEDED AND CAN SUCCEED.

When we think of critical thinking, we think of something *bigger* than its domain of training. When I teach students how to evaluate the argument in a set of newspaper editorials, I am hoping that they will learn to evaluate arguments generally, not just those they read, and not just those they would find in other editorials. This aspect of critical thinking is called transfer, and the research literature evaluating how well critical thinking skills transfer to new problems is decidedly mixed.

TEACHING CRITICAL THINKING FOR GENERAL TRANSFER

It is self-evident that we expect *some* transfer in learning. An extreme version of transfer failure might be, for example, the inability to graph any functions except the exact same ones graphed in class. We could take transfer to the other extreme and propose perfect general transfer, meaning that mental work prompts improvement in *any other* mental work, no matter how far removed; for example, learning to graph linear functions makes one a better writer. Improbable as it seems, this idea has been taken seriously for many years.

The earliest and likely most enduring version was termed formal discipline, the idea that studying difficult content trained a student's will and perhaps attention; difficult work taught students to focus and stick to a task. In addition, advocates suggested that some subjects—Latin, for example, or geometry—demanded logical thinking, which would prompt students to think logically in other contexts (Lewis, 1905).

The idea was challenged by psychologist Edward Thorndike, whose theory of human learning suggested that such transfer was impossible. Thorndike conducted a series of experiments showing that practice on one task (estimating the areas of rectangles) did not yield a benefit to other seemingly similar tasks, like estimating the area of other geometric shapes (Thorndike & Woodworth, 1901). Thorndike conducted a more pointed test of the formal discipline idea two decades later (Broyler, Thorndike, & Woodyard, 1924; Thorndike, 1923). High school students took standardised tests in autumn and spring, and Thorndike analyzed the difference in scores for each student as a function of the coursework they had taken during the year. If Latin, for example, makes you smart, students who take it should score better in the spring. The results did not support formal discipline.

But the theory did not die. For one thing, Thorndike's methods were open to criticism (see Rosenblatt, 1967). More importantly, a new task emerged that seemed a better bet to teach logical thinking: computer programming. In the 1960s computer scientist Seymour Papert led calls for young students to learn computer programming, with the idea that doing so would improve their thinking abilities (Papert, 1972, 1980; see also Clements & Gullo, 1984; Linn, 1985). Studies through the 1980s showed mixed results (Liao & Bright, 1991) but calls were renewed in the early 21st century, as the need for computational thinking in the emerging job market seemed more urgent than ever (Grover & Pea, 2013; Wing, 2008).

A recent meta-analysis offers some apparently encouraging results about the general trainability of computational thinking (Scherer, Siddiq, & Viveros, 2018). The researchers reported that learning to program a computer yields positive transfer to measures of creative thinking, mathematics, metacognition, spatial skills, and reasoning, with an average effect size of g = .47.[1] The authors note that effects were considerably smaller when studies used an active control group (that is, students who didn't learn to program undertook some other

[1] Hedge's g is a measure of effect size, very similar to Cohen's d—it includes a correction for bias in small samples that Cohen's d does not include. An effect size of g = .47 would conventionally be considered of medium size.

IT IS NO SURPRISE THAT PROGRAMS IN SCHOOL MEANT TO TEACH GENERAL CRITICAL THINKING SKILLS HAVE HAD LIMITED SUCCESS.

special activity) which may indicate a placebo effect accounts for at least part of the benefit. Of what remains, it is sensible to think that this transfer was a consequence of conceptual overlap between programming and these skills, as no benefit was observed in measures of literacy.

Hopeful adults have tried still other activities as potential all-purpose enhancers of intelligence, for example exposure to classical music, the so-called Mozart effect (Pietschnig, Voracek, & Formann, 2010), learning to play a musical instrument (Sala & Gobet, 2017), or learning to play chess (Sala & Gobet, 2016). None have succeeded as hoped.

It is no surprise then that programs in school meant to teach general critical thinking skills have had limited success. Such programs are usually curricular add-ons during which students engage in critical thinking activities for perhaps five hours each week over the course of a year or two. Unfortunately, the evaluations of these programs seldom offer a rigorous test of transfer. If the critical thinking training features logical and spatial puzzles, the measure of success tends to feature the same sort of puzzle (see Kozulin et al., 2010). And if the critical thinking regimen entails argument and debate, the outcome measure is usually the ability to evaluate arguments or take both perspectives in debate (see Kuhn & Crowell, 2011; Reznitskaya et al., 2012). When investigators have tested for transfer in such curricular programs, positive results have been absent or modest and quick to fade (Ritchart & Perkins, 2005).

TRANSFER AND THE NATURE OF CRITICAL THINKING

What do these results tell us about the nature of critical thinking? They tell us something that we perhaps should have recognised with a bit of reflection. It is not useful to think of critical thinking skills, once acquired, as broadly applicable. Wanting students to be able to "analyse, synthesise and evaluate" information sounds like a reasonable goal. But analysis, synthesis, and evaluation mean different things in different disciplines. Literary criticism has its own internal logic, its norms for what constitutes good evidence and a valid argument. These norms differ from those found in mathematics, for example. And indeed, different domains—science and history, say—have different definitions of what it means to "know" something. Thus, our goals for student critical thinking must be domain-specific. An overarching principle like "think logically" is not a useful goal.

But wait. Surely there are *some* principles of thinking that apply across fields of study. "A" and "not A" cannot be simultaneously true, whether the domain is mathematics or history. Denial of the consequent is always wrong, strawperson arguments are always weak, and having a conflict of interest always makes your argument suspect (Ennis, 1987). There are indeed principles that carry across domains of study. The problem is that people who learn these broadly applicable principles in one situation often fail to apply them in a new situation.

The law of large numbers provides an example. It states that a large sample will probably be closer to a "true" estimate than a small sample—if you want to know whether a set of dice is loaded, you're better off seeing the results of 20 throws rather than two throws. People readily understand this idea in the context of evaluating randomness, but they are less likely to see the need for a large sample when judging academic performance; they are ready to say that someone who received poor grades on two maths tests is simply bad at maths. And

IT IS NOT USEFUL TO THINK OF CRITICAL THINKING SKILLS, ONCE ACQUIRED, AS BROADLY APPLICABLE ... ANALYSIS, SYNTHESIS AND EVALUATION MEAN DIFFERENT THINGS IN DIFFERENT DISCIPLINES.

when it comes to social behaviour, they are even less interested in the law of large numbers; they think that a brief observation of a person's friendliness tells you whether or not that person is friendly (Jepson, Krantz, & Nisbett, 1983).

This surprising failure to deploy useful knowledge can even be observed immediately after learning. In a classic experiment, researchers administered a tricky problem. Subjects were told that a malignant tumour could be treated with a particular type of ray, but that the ray caused a lot of collateral damage to healthy tissue. How, subjects were asked, could the ray be used to destroy the tumour? Few were able to solve the problem in the allotted 20 minutes. Other subjects read a story describing a military situation analogous to the medical problem. Instead of rays attacking a tumour, rebels were to attack a dictator hiding out in a fortress. The military story described the solution, but despite reading it moments before they tried the medical problem, subjects didn't see the analogy: disperse the forces to avoid collateral damage and have forces converge at the point of attack. Merely mentioning that the story might help solve the problem boosted solution rates to nearly 100 per cent. Thus, using the analogy was not hard; the problem was thinking to use it in the first place (Gick & Holyoak, 1980, 1983).

These results offer a new perspective on critical thinking. The problem in transfer is not just that different domains have different norms for critical thinking. The problem is that previous critical thinking successes seem encapsulated in memory. We know that a student has understood an idea like the law of large numbers. But understanding it offers no guarantee that the student will recognise new situations in which that idea will be useful.

CRITICAL THINKING AS PROBLEM RECOGNITION

We seem to face a significant challenge: how can we improve student critical thinking if it is difficult for them to appreciate that some new problems are actually ones they have solved in the past? Let us elaborate on the challenge that thinkers face in transferring knowledge. The rays/tumour problem shares what researchers call "deep structure" with the rebels/fortress problem. The deep structure is *distribute forces to avoid inflicting collateral damage, and converge at the point of attack*. The "surface structures" of the problems differ: in one case the problem appears to be about rays and tumours and the other is about rebels and a fortress. A key challenge to thinking critically is that, when confronted with a problem, we tend to dwell on the surface structure and so we fail to realise we've thought through a problem before.

Why? Probably because the surface structure is explicit, obvious. Who could be blamed for searching memory for information related to doctors, hospitals, tumours, and so on, given that's what the problem is about? And just as obviously, the deep structure is not explicit. There's no overt hint that the deep structure is *distribute forces to avoid inflicting collateral damage, and converge at the point of attack*. For all the reader knows the deep structure is *don't make a hasty generalisation, or correlation is not causation, or Newton's third law of motion*. Deep structure is defined by functional relationships

A KEY CHALLENGE TO THINKING CRITICALLY IS THAT, WHEN CONFRONTED WITH A PROBLEM, WE TEND TO DWELL ON THE SURFACE STRUCTURE AND SO WE FAIL TO REALISE WE'VE THOUGHT THROUGH A PROBLEM BEFORE.

among the elements of the described situation, and there are many possible functional relationships.

You might think, then, that the sensible approach would be to teach deep structure in the first place. If people get stuck on surface structure, why not avoid that altogether? The problem is that deep structure is usually abstract and difficult to understand. If a teacher simply said "For every action there is an equal and opposite reaction" students would not understand the principle, and would ask for examples. In other words, they would demand descriptions with rich surface structure.

Happily, this difficulty in recognising problems does disappear in the face of significant practice. Experts report that they "just see" the structure of complex problems in their domain of expertise, and even non-experts "just see" solutions to problems they have seen many times. For example, my middle-schooler pointed out an article on social media, thinking it was news reporting. I immediately recognised that was an op-ed. My daughter has studied text genres at school, yet did not see the connection. Why do I see it whereas she does not? It's reasonable to suppose that the important feature is seeing a large number of problems with varying surface structures but the same deep structure (see Chen & Mo, 2004). That is fine if you are not in a hurry, but is it possible to hasten the recognition of deep structure, and thus boost transfer?

As noted, deep structure is so abstract that novices find it difficult to understand, but there are ways of getting students to think about deep structure even as you present the problem with rich surface structure. One technique is problem comparison: show students two solved problems with different surface structures but the same deep structure and ask them to compare them (Kurtz, Boukrina, & Gentner, 2013). The similarities obviously occur at the deep level, so the process of comparison prompts thinking about that deep structure. In one experiment testing this method, business school students were asked to compare two stories, one involving international companies coping with a shipping problem, and the other concerning two college students planning a spring break trip. In each, a difficult negotiation problem was resolved through the use of a particular type of contract. Two weeks later, students were more likely to use the solution on a novel problem if they had contrasted the stories compared to other students who simply read them (Loewenstein, Thompson, & Gentner, 1999).

Richard Catrambone developed a different technique to address a slightly different transfer problem. He noted that in maths and science classes, students often learned to solve standard problems via a series of fixed, lock-step procedures. That meant students were stumped when confronted with a problem requiring a slight revision of the steps, even if the goal of the steps was the same. For example, a student might learn a method for solving work problems like "Trisha can paint a house in 14 hours and Carole can do it in eight. How long would it take them to paint one house, working together?" A student who learns a sequence of steps to solve that sort of problem is often thrown by a small change—

the homeowner had already painted a quarter of the house before hiring Trisha and Carole.

Catrambone (Catrambone, 1995, 1998; Catrambone & Holyoak, 1990; Margulieux & Catrambone, 2016) showed that student knowledge will be more flexible if students are taught to label the substeps of the solution with the goal it serves. For example, work problems are typically solved by calculating how much of the job each worker can do in an hour. If, during learning, that step were labelled so students understood that that calculation was part of deriving the solution, they would know how to solve the problem when a fraction of the house is to be painted.

OPEN-ENDED PROBLEMS AND KNOWLEDGE

Students encounter standard problems that are best solved in a particular way, but many critical thinking situations are unique. Critical thinking is needed when playing chess, designing a product, or planning strategy for a field hockey match. But there are no routine, reusable solutions for these problems. Nevertheless, just as with routine problems, critical thinking for open-ended problems is enabled by extensive stores of knowledge about the domain (North et al., 2011). Knowledge aids critical thinking in three ways.

First, the recognition process described above ("oh, this is that sort of problem") can still apply to subparts of a complex, open-ended problem. Complex critical thinking may entail multiple simpler solutions from memory that can be "snapped together" when solving complex problems (Koedinger, Corbett, & Perfetti, 2012; Taatgen, 2013). In an intuitive example, calculating the best value among several vacation packages may be a novel, open-ended problem, but if the method of comparison calls for long division, I don't need to think through a method to execute that substep.

Furthermore, the process of recognition may not tell the thinker exactly what to do, but it may help the thinker evaluate the situation more quickly or accurately. For example, classic work on expertise in chess showed that expert players use recognition processes to evaluate chess positions (Chase & Simon, 1973). Experts have literally thousands of mid-play board positions in memory, and comparing current play to those positions helps chess experts identify which parts of their position (and their opponent's) is strong or weak. That, in turn, helps them determine where to focus their attention as they plan their next move.

The second way that knowledge contributes to critical thinking in open-ended problems is through its impact on working memory. Working memory refers, colloquially, to the place in the mind where thinking happens—it's where you hold information and manipulate it to carry out cognitive tasks. So for example if I said "how is a scarecrow like a blueberry?" you would retrieve information about scarecrows (not alive, protect crops, found in fields, birds think they are alive) and blueberries (purple, used in pies, small, featured in *Blueberries for Sal*) from your memory and then you'd start comparing these features, looking for overlap.

JUST AS WITH ROUTINE PROBLEMS, CRITICAL THINKING FOR OPEN-ENDED PROBLEMS IS ENABLED BY EXTENSIVE STORES OF KNOWLEDGE ABOUT THE DOMAIN.

An important feature of working memory is its limited size. Suppose I said "What do these objects have in common: a blueberry, a scarecrow, a flowerpot, a drumstick, and a dishwasher?" Working memory would be overwhelmed. There's probably space for the five words, but not for the five words, *and* a bunch of information about each word, *and* left over attention to compare them. So how does a chess player think about all 32 pieces on the board and their relative positions (to assess the current game), *and* have attention left over to contemplate effective moves?

We've already seen one way that knowledge helps—recognising the current board position as similar to a previously seen board helps the player recognise areas of strength and weakness. In addition, knowledge allows the player to treat groups of pieces as a single unit. The king, a castle and three pawns in a corner of the board relate to one another in the defensive position, so the expert will treat them as a single unit. This ability to clump multiple entities into a single, meaningful unit has been observed in many domains of expertise, as varied as dance, circuit design and computer programming. When experience allows you to unite many separate dance moves into a single unit, it saves working memory space. That allows more working memory space for the dancer to think about more subtle aspects of movement, rather than crowding working memory with "what I am to do next."

The third way that knowledge may contribute to critical thinking is in enabling you to deploy thinking strategies. When we discussed the recognition of deep structure, the problem was that you had an effective thinking strategy in your memory but you failed to retrieve it because you did not see that it was relevant. But some situations that call for critical thinking are easily labelled and recognised. We can tell students that they should evaluate the logic of the author's argument when they read an op-ed, and we can tell them the right method to use when conducting a scientific experiment. Students should have no trouble recognising "Oh, this is that sort of problem" and they may have committed to memory the right thinking strategy. They know what to do, but they may not be able to use the strategy without the right domain knowledge.

This point is rather obvious in the case of a critical thinking skill like evaluating an argument: abstract principles like "look for hidden assumptions" won't help much in sizing up an op-ed about the war in Afghanistan if you know very little about the topic. Never mind evaluating the argument in the op-ed, if you lack background knowledge about the topic, ample evidence from the last 40 years indicates you will not comprehend the author's claims in the first place (Willingham, 2017). That is because writers (and speakers) omit information they assume their audience already knows. For example, a writer might warn that the US could "find itself in the Soviet role in this long-standing war," assuming that the reader knows that the Soviet Union fought a costly, unsuccessful war there in the 1980s.

The importance of background knowledge to critical thinking extends beyond reading. Principles of scientific reasoning seem to be content free: "a control group should be identical to the experimental group, except for the treatment," for example, or "theories should be as simple as possible, while accounting for all the data." In practice, however, content knowledge is needed to use the principles. For example in an experiment on learning, you would want to be sure that the experimental and control groups were comparable, so you would make sure that, for example, proportions of men and women in each group were the same. What characteristics besides sex should you be sure are equivalent in the experimental and control groups? Ability to concentrate? Intelligence? You can't measure every characteristic of your subjects, so you

EXPERIMENTAL EVIDENCE SHOWS THAT AN EXPERT DOES NOT THINK AS WELL OUTSIDE HER AREA OF EXPERTISE, EVEN IN A CLOSELY RELATED DOMAIN. SHE IS STILL BETTER THAN A NOVICE, BUT HER SKILLS DO NOT TRANSFER COMPLETELY.

would focus on characteristics that you know are relevant to learning. But knowing which characteristics are "relevant to learning" means knowing the research literature in learning and memory. It is simple to define "control group" (and simple for students to memorise the definition and repeat it on a test) but using the definition to create a good control group depends on knowledge of the domain under study.

Experimental evidence shows that an expert does not think as well outside her area of expertise, even in a closely related domain. She is still better than a novice, but her skills do not transfer completely. For example, knowledge of medicine transfers poorly among subspecialties; neurologists do not diagnose cardiac cases well (Rikers, Schmidt, & Boshuizen, 2002). Expertise in writing is similarly encapsulated; a technical writer who specialises in writing instruction pamphlets for home electronics can't write newspaper articles (Kellogg, 2018). Perhaps most surprisingly the analytic abilities of professional philosophers do not extend to everyday judgments. Philosophers are no less susceptible than average adults to being swayed by irrelevant features of problems like question order or wording (Schwitzgebel & Cushman, 2015).

HOW TO TEACH STUDENTS TO THINK CRITICALLY

So what does all this mean? Is there really no such thing as a "critical thinking skill" if by "skill" we mean something generalisable? Is everything that we might be tempted to call a skill actually hyper-specific?

Maybe, but it is hard to be sure. It is certainly the case that psychologists have had a difficult time proving the utility of general-purpose thinking skills, things like applying rules of deductive logic. We do know that students who go to school longer score better on intelligence tests, and certainly we think of intelligence as all-purpose (Carlsson et al., 2015; Ritchie & Tucker-Drob, 2018; Strenze, 2007). One interpretation is that people score better because they have learned a lot of fairly specific thinking skills, and there is, therefore, a higher probability that the intelligence test probes something they know about. It remains possible that more and/or better schooling yields an advantage via thinking skills that are more general, and researchers are simply unable to identify them, but existing data favour the specific skills account (Ritchie, Bates, & Deary, 2015).

If you are a researcher, it is unclear which view of the good critical thinker to adopt: someone who has mastered lots of specific skills, or someone with a smaller set of yet-to-be-identified general skills. But educators are not researchers, and for educators, one fact ought to be salient. We are not even sure the general skills exist, but we are quite sure there is no proven way to teach them directly. In contrast, we have a pretty good idea of how to teach students the more specific critical thinking skills. I suggest we do so. Here is a four-step plan.

First, identify what is meant by critical thinking in each domain. Be specific. What tasks showing critical thinking should a high school

graduate be able to do in mathematics, history, and other subjects? It is not useful to set a goal that students "think like historians," or "learn the controversies surrounding historical events." If students are to read as historians do, they need to learn specific skills like interpreting documents in light of their sources, corroborating them, and putting them in historical context. Notably, skilful reading is different in other disciplines. Scientists believe that the source of a document is irrelevant so long as it is trustworthy. And unlike historical documents, scientific documents are written in a consistent format. Learning to read like a scientist means, in part, learning the conventions of this format.

These skills should be explicitly taught and practiced—there is evidence that simple exposure to this sort of work without explicit instruction is less effective (Abrami et al., 2008; Halpern, 1998; Heijltjes, Van Gog, & Paas, 2014). In addition, it is clear that educators will have to pick and choose which skills their students will learn. Even across the long thirteen years of schooling, time is limited.

Second, identify the domain content that students must know. We have seen that domain knowledge is a crucial driver of thinking skills. For example, sourcing historical documents means interpreting their content in light of the author, the intended audience, and circumstances under which the author wrote. It is not enough to know that a letter was written by an army sergeant to his wife just before the Battle of Romani. The student must know enough about the historical context to understand how this sourcing information ought to influence his or her interpretation of the letter.

What knowledge is essential to the type of thinking you want your students to be able to do? That of course depends on one's educational goals. We might suggest that students should focus on content that is most likely to lead to rewarding employment, or, given that schools are publicly supported, content that enables them to become active and informed citizens and inspiring leaders. Or perhaps the purpose of school is to help students better understand their individual abilities and passions, and content should accordingly be personalised.

The prospect of someone deciding which knowledge students ought to learn—and what they won't learn—sometimes makes people uneasy exactly because this decision depends on one's goals for schooling, and goals depend on values. Selection of content is a critical way that values are expressed (Willingham, 2012). Making that choice will lead to uncomfortable trade-offs. But not choosing is still making a choice. It is choosing not to plan, and to let random forces determine what students learn.In the third step, educators must select the best sequence in which to learn the skills. It is obvious that skills and knowledge build on one another in mathematics, or in history; it is easier to understand why the Constitution of Australia was approved at the turn of the 20th century if one knows why Australians were concerned about the French and German presence in the South Pacific in the late 19th century. What is true of maths and history is true of other domains of skill and knowledge; we interpret new information in light of what we already know. The right preparation makes new learning easier.

Fourth, educators must decide which skills should be revisited across years. Studies show that even if content is learned quite well over the course of half of a school year, about half will be forgotten in three years (Pawl et al., 2012). That doesn't mean there's no value in exposing students to content just once; most students will forget much but they'll remember something, and for some students, an interest may be kindled. But when considering skills we hope will stick with students for the long term, we should plan on at least three to five years of practice (Bahrick, 1984; Bahrick & Hall, 1991). Most of the time, this practice will look different—it will be embedded in new skills and content. But this revisiting should be assured and planned.

SOME PRACTICAL MATTERS OF TEACHING CRITICAL THINKING

I have outlined a broad, four-step plan. Let us consider some of the pragmatic decisions educators face as they contemplate the teaching of critical thinking.

Is it all or none? I have suggested that critical thinking be taught in the context of a comprehensive curriculum. Does that mean an individual teacher cannot do anything on his or her own? Is there just no point in trying if the cooperation of the entire school system is not assured?

Obviously that is not the case; a teacher can still include critical thinking content in his or her courses and students will learn, but it is quite likely they will learn more, and learn more quickly if their learning is coordinated across years. It has long been recognised among psychologists that an important factor, perhaps the most important factor, influencing learning is what the student already knows (Ausubel, 1968). Teaching will be more effective if the instructor is confident about what his or her students already know.

Student age: When should critical thinking instruction start? There is not a firm, research-based answer to this question. Researchers interested in thinking skills like problem solving, or evidence evaluation in young children (preschool through early primary school), have studied how children think in the absence of explicit instruction. They have not studied whether or how young children can be made to think more critically. Still, research over the last thirty years or so has led to an important conclusion: children are more capable than we thought.

The great developmental psychologist Jean Piaget proposed a highly influential theory that suggested children's cognition moves through a series of four stages, characterised by more abstract thought, and better ability to take multiple perspectives. In

RESEARCH OVER THE LAST THIRTY YEARS OR SO HAS LED TO AN IMPORTANT CONCLUSION: CHILDREN ARE MORE CAPABLE THAN WE THOUGHT.

stage theories, the basic architecture of thought is unchanged for long periods of time, and then rapidly reorganises as the child moves from one developmental stage to another (Piaget, 1952). A key educational implication is that it is at least pointless and possibly damaging to ask the child to do cognitive work that is appropriate for a later developmental stage. The last thirty years has shown that, contrary to Piaget's theory, development is gradual and does not change abruptly. It has also shown that what children can and cannot do varies depending on the content. For example, in some circumstances, even toddlers can understand principles of conditional reasoning, and in other circumstances, conditional reasoning confuses adult physicians. It all depends on the content of the problem (Willingham, 2008).

Thus, research tells us that including critical thinking in the schooling of young children is likely to be perfectly appropriate. It does not, however, provide guidance into what types of critical thinking skills to start with. That is a matter to take up with experienced educators, coordinating with colleagues who teach older children in the interests of making the curriculum seamless.

Types of students: Should everyone learn critical thinking skills? The question sounds like a set-up, like an excuse for a resounding endorsement of critical thinking for all. But the truth is that, in many systems, less capable students are steered into

less challenging coursework, with the hope that by reducing expectations, they will at least achieve mastery of the basics, or these lower expectations may pervade entire schools that serve students from low-income families (Parker et al., 2016).

It is worth highlighting that access to challenging content and continuing to tertiary education is, in nearly every country, associated with socio-economic status (OECD, 2018). Children from high socio-economic status families also have more opportunities to learn at home. If school is the chief or only venue through which low socio-economic status students are exposed to advanced vocabulary, rich content knowledge, and demands for high-level thinking, it is absolutely vital that those opportunities be enhanced, not reduced.

Assessment: Assessment of critical thinking is, needless to say, a challenge. One difficulty is expense. Claims to the contrary, multiple choice items do not necessarily require critical thinking, even when items are carefully constructed and vetted, as on the United States National Assessment of Educational Progress (NAEP). Smith (2017) administered items from the history NAEP for 12th graders to college students who had done well on other standardised history exams. Students were asked to think aloud as they chose their answers, and the researchers observed little critical thinking, but a lot of "gaming" of the questions. Assessing critical thinking requires that students answer open-form questions, and that means humans must score the response, an expensive proposition.

On the bright side, the plan for teaching critical thinking that I have recommended makes some aspects of assessment more straightforward. If the skills that constitute critical thinking in, say, 10th grade chemistry class are fully defined, then there is no question as to what content ought to appear on the assessment. This predictability ought to make teachers more confident that they can prepare their students for standardised assessments.

As much as teaching students to think critically is a universal goal of schooling, one might be surprised that student difficulty in this area is such a common complaint. Educators are often frustrated that student thinking seems shallow. This review should offer insight into why that is. The way the mind works, shallow is what you get first. Deep, critical thinking is hard-won.

That means that designers and administrators of a program to improve critical thinking among students must take the long view, both in the time frame over which the program operates, and especially the speed with which one expects to see results. Patience will be a key ingredient in any program that succeeds.

REFERENCES

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research*, 78(4), 1102–1134.

Bensley, D. A., & Spero, R. A. (2014). Improving critical thinking skills and metacognitive monitoring through direct infusion. *Thinking Skills and Creativity*, 12, 55–68.

Ausubel, D. P. (1968). *Educational Psychology*: A Cognitive View. New York: Holt, Rinehart & Winston.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology:* General, 113(1), 1–29.

Bahrick, H. P., & Hall, L. K. (1991). Lifetime maintenance of high school mathematics content. *Journal of Experimental Psychology: General*, 120(1), 20–33.

Bangert-Drowns, R. L., & Bankert, E. (1990). Meta-analysis of effects of explicit instruction for critical thinking. In A*nnual Meeting of the American Educational Research Association* (pp. 56–79). Boston, MA.

Broyler, C. R., Thorndike, E. L., & Woodward, E. (1924). A Second Study of Mental Discipline in High School Studies. *Journal of Educational Psychology*, 18(6), 377–404.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119(2), 197–253.

Carlsson, M., Dahl, G. B., Öckert, B., & Rooth, D.-O. (2015). The Effect of Schooling on Cognitive Skills. *Review of Economics and Statistics*, 97(3), 533–547.

Catrambone, R. (1995). Aiding Subgoal Learning: Effects on Transfer. *Journal of Educational Psychology*, 87(1), 5–17.

Catrambone, R. (1998). The subgoal learning model: Creating better examples to improve transfer to novel problems. Journal of Experimental Psychology: General, 127(4), 355–376.

Catrambone, R., & Holyoak, K. J. (1990). Learning subgoals and methods for solving probability problems. *Memory & Cognition*, 18(6), 593–603.

Chase, W. G., & Simon, H. A. (1973). The Mind's Eye in Chess. In W. G. Chase (Ed.), *Visual Information Processing* (pp. 215–281). New York: Academic Press.

Chen, Z., & Mo, L. (2004). Schema induction in problem solving: A multidimensional analysis. *Journal of Experimental Psychology: Learning Memory and Cognition*, 30(3), 583–600.

Clements, D. H., & Gullo, D. F. (1984). Effects of computer programming on young children's cognitions. *Journal of Educational Psychology1*, 76, 1051–1058.

Ennis, R. H. (1987). Critical Thinking and the Curriculum. In *Thinking Skills Instruction : Concepts and Techniques* (pp. 40–48).

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. Cognitive Psychology, 12(3), 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1–38.

Grover, S., & Pea, R. (2013). Computational Thinking in K-12: A Review of the State of the Field. *Educational Researcher*, 42(1), 38–43.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4), 449–455.

Heijltjes, A., Van Gog, T., & Paas, F. (2014). Improving students' critical thinking: Empirical support for explicit instructions combined with practice. *Applied Cognitive Psychology*, 28(4), 518–530.

Helsdingen, A. S., van den Bosch, K., van Gog, T., & van merrienboer, J. J. G. (2010). The Effects of Critical Thinking Instruction on Training Complex Decision Making. *Human Factors*, 52(4), 537–545.

Jepson, C., Krantz, D. H., & Nisbett, R. E. (1983). Inductive reasoning: Competence or skill? *Behavioral and Brain Sciences*, 6(03), 494.

Kellogg, R. T. (2018). Professional writing expertise. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge, UK: Cambridge University Press.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction Framework: Bridging the Science-Practice Chasm to Enhance Robust Student Learning. *Cognitive Science*, 36(5), 757–798.

Kozulin, A., Lebeer, J., Madella-Noja, A., Gonzalez, F., Jeffrey, I., Rosenthal, N., & Koslowsky, M. (2010). Cognitive modifiability of children with developmental disabilities: A multicentre study using Feuerstein's Instrumental Enrichment-Basic program. *Research in Developmental Disabilities*, 31(2), 551–559.

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, 22(4), 545–552.

Kurtz, K. J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *Journal of Experimental Psychology: Learning Memory and Cognition, 39*(4), 1303–1310.

Lewis, F. C. (1905). A study in formal discipline. *The School Review*, 13(4), 281–292.

Liao, Y.-K. C., & Bright, G. W. (1991). Effects of Computer Programming on Cognitive Outcomes: A Meta-Analysis. *Journal of Educational Computing Research*, 7(3), 251–268.

Linn, M. C. (1985). The cognitive consequences of programming instruction in classrooms. *Educational Researcher*, 14(5), 14–29.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin and Review*, 6(4), 586–597.

Margulieux, L. E., & Catrambone, R. (2016). Improving problem solving with subgoal labels in expository text and worked examples. *Learning and Instruction*, 42, 58–71.

North, J. S., Ward, P., Ericsson, A., & Williams, A. M. (2011). Mechanisms underlying skilled anticipation and recognition in a dynamic and temporally constrained domain. *Memory*, 19(2), 155–168.

OECD (2018). *Education at a Glance 2018: OECD Indicators.* OECD Publishing: Paris.

Papert, S. (1972). Teaching Children to be Mathematicians Versus Teaching About Mathematics. *International Journal of Mathematical Education in Science and Technology,* 3(3), 249–262.

Papert, S. (1980). Mindstorms. New York: Basic Books.

Parker, P. D., Jerrim, J., Schoon, I., & Marsh, H. W. (2016). A multination study of socioeconomic inequality in epectations for progression to higher education: The role of between-school tracking and ability stratification. *American Educational Research Journal,* 53 (1), 6-32.

Pawl, A., Barrantes, A., Pritchard, D. E., & Mitchell, R. (2012). What Do Seniors Remember from Freshman Physics?, Physical *Review Special Topics—Physics Education Research 8*, 020118.

Piaget, J. 1952. *The Origins of Intelligence in Children.* New York: International Universities Press.

Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Mozart effect-Shmozart effect: A meta-analysis. *Intelligence*, 38(3), 314–323.

Reznitskaya, A., Glina, M., Carolan, B., Michaud, O., Rogers, J., & Sequeira, L. (2012). Examining transfer effects from dialogic discussions to new tasks and contexts. *Contemporary Educational Psychology*, 37(4), 288–306.

Rikers, R. M. J. P., Schmidt, H. G., & Boshuizen, H. P. A. (2002). On the constraints of encapsulated knowledge: Clinical case representations by medical experts and subexperts. *Cognition and Instruction,* 20(1), 27–45.

Ritchart, R., & Perkins, D. N. (2005). Learning to think: The challenges of teaching thinking. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 775–802). Cambridge, UK: Cambridge University Press.

Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental Psychology*, 51(5), 573–582.

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How Much Does Education Improve Intelligence? *A Meta-Analysis. Psychological Science*, 29(8), 1358–1369.

Sala, G., & Gobet, F. (2016). Do the benefits of chess instruction transfer to academic and cognitive skills? A meta-analysis. *Educational Research Review*, 18, 46–57.

Sala, G., & Gobet, F. (2017). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review*, 20, 55–67.

Scherer, R., Siddiq, F., & Viveros, B. S. (2018). The Cognitive Benefits of Learning Computer Programming: A Meta-Analysis of Transfer Effects. *Journal of Educational Psychology*.

Schwitzgebel, E., & Cushman, F. (2015). Philosophers' biased judgments persist despite training, expertise and reflection. *Cognition*, 141, 127–137.

Smith, M. D. (2017). Cognitive validity: Can multiple-choice items tap historical thinking processes? *American Educational Research Journal*, 54, 1256–1287.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, 35(5), 401–426.

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review*, 120(3), 439–471.

Thorndike, E. L. (1923). The influence of first-year latin upon ability to read English. *School and Society*, 17, 165–168.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upone the efficiency of other functions (I). *1Psychological Review*, 8, 247–261.

Willingham, D. T. (2008). What is developmentally appropriate practice? *American Educator, Summer,* 34-39.

Willingham, D. T. (2012). *When can you trust the experts? How to tell good science from bad in education.* San Francisco, CA: Jossey-Bass.

Willingham, D. T. (2017). The reading mind. San Francisco, CA: Jossey-Bass.

Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 366*(1881), 3717–3725.

© May 2019.

Education: Future Frontiers | Occasional Paper Series