Since the 1960s, concern that American students may not be capable of transferring the skills they have gained from their education to the practical problems of life has troubled educators. Of greatest concern is whether students have mastered “critical thinking” or “higher order thinking skills” and can apply them outside of school curricula. These concerns have given rise to the “critical thinking movement.”

To demonstrate that the movement is successful, it must prove that its efforts not only increase the critical thinking of students in school, but that students can transfer critical thinking to novel situations, including those encountered in daily life. The primary purpose of this review is to ascertain if there is compelling evidence that efforts to teach critical thinking have had this result.

What became apparent in the process of this review, however, was that several subsidiary problems must first be answered before the problem of evaluating the effectiveness of critical thinking transfer can be approached. The first of these problems is whether the movement has a common theme or definition of “critical thinking.” Second is the question, does “critical thinking” encompass “creative thinking” or is it antithetical to it. The third problem might be formulated thus: is “critical thinking” generalizable or is it tied to subject matter? The fourth problem is whether adequate evaluative measures of critical thinking are available to measure the effectiveness of efforts to teach critical thinking. Answering these prior question is essential before inquiring whether there is compelling evidence that teaching critical thinking results in a transfer of skills or dispositions that students can use in other arenas. This line of inquiry supplies the structure for this review of the relevant literature.

The scope of this review is limited. Most critical thinking literature provides program and instructional technique description. This material is out of scope for this review except as it bears directly upon the question concerning subject-dependence in relation to critical thinking. Furthermore, although this review addresses the works of most seminal thinkers in the critical thinking movement, constraints and limited access to information means that some major figures, such as Harvey Siegel, have not been included. Finally, although philosophical literature on this subject abounds, evaluative studies using either qualitative or quantitative methods to measure the effectiveness of whole programs are comparatively scarce. I have included relevant examples of these studies, yet it can be said at the outset that the dearth such studies needs to be redressed by the research community.

The Common Theme of the Critical Thinking Movement

The first step in understanding the Critical Thinking Movement is to uncover the essential characteristics of critical thought and examine the commonality of agendas for
Proponents of the critical thinking movement posit numerous reasons for teaching critical thinking. A common reason is a reflection of the shift in economic patterns away from an industrial society into arenas in which laborers must solve complex problems (Bloom; Reich; Paul; Nickerson). Another reason frequently proposed is that critical thinking skills are necessary for effective citizenship in a democracy, for example, in selecting leaders and being a juror (Ennis, “Taxonomy”; Paul; Nickerson). Paul and Nickerson also call attention to the capacity of human beings for self-delusion and note that irrational human behavior causes great suffering in the world. They see critical thinking the antidote. Finally, both these thinkers uphold the notion that thinking is a significant part of being human; therefore, mastery of critical thinking is a necessary for being a fully developed human being.

Another premise of the proponents of the critical thinking movement is that critical thinking does not always unfold naturally as a part of growth. Furthermore, critical thinking is not effectively taught in traditional school settings that rely heavily upon rote memorization and didactic teaching methods (Kennedy; Paul; Nickerson; Schrag). Therefore, leaders of the movement have developed numerous programs to teach critical thinking.

The common theme of the critical thinking movement is that critical thinking skills involve the ability to make reasonable decisions in complex situations, such as those found in a rapidly changing and complex society. The movement emphasizes “knowing how” more than “knowing that” (Roland). Furthermore, helping individuals gain these abilities, requires a self-conscious attempt on the part of educators to address the cultivation of critical thinking by utilizing methods other than simply rote memorization and didactic instruction.

What is Critical Thinking?

The unity of the movement disintegrates once the question “what do you mean by critical thinking?” is asked. There is a significant divergence of opinion about what constitutes critical thinking.

Some scholars identify critical thinking with the mastery of specific skill sets and provide schematics or taxonomies to express their inter-relationships. A Committee of College and University Examiners created one of the early taxonomies (Bloom). Bloom and his colleagues identified six major classes of cognitive skills: Knowledge (by which they mean recall); Comprehension; Application; Analysis; Synthesis; and Evaluation. One reason for this construction is that the lower skills are required in order for the higher skills to be used. Comprehension requires Knowledge, or recall. Therefore, critical thinking, in Bloom’s view, is gaining mastery of these skill sets and selecting the appropriate techniques when encountering a novel situation.

The primary strength of Bloom’s taxonomy is that it is logical and hierarchical, guiding the educator in a process leading from the most simple to the most complex form of cognitive skills. It is also comparatively easy to evaluate the mastery of these skills because they link to particular behaviors (Bloom 12). Bloom supplies numerous evaluative techniques linked to the taxonomy.
There are, however, disadvantages with Bloom’s taxonomy. Historically, many teachers have used it as a “cookbook” without demonstrating critical thinking skills themselves (Paul 375-383). Paul also criticizes Bloom for overemphasizing recall and for insisting on neutrality. Paul believes that critical thinking should be used to reach substantial value judgments. Finally, Paul conceives Bloom’s taxonomy as neglecting the dialectical dynamic of critical thinking. With regard to his first point, Paul overstates his case; misuse of the taxonomy does not invalidate the design itself. The emphasis given to Knowledge or recall is more controversial, relating to the question whether or not critical thinking is subject-dependent. I do believe that Paul is correct in criticizing Bloom’s view that critical thinking is value neutral, since real life decisions are never value neutral; but again, that does not invalidate the structure of his taxonomy. The neglect of the dialectical process in critical thinking, however, is a substantial criticism that does seem borne out in the construction of the taxonomy, which is designed to flow from simple to complex.

Another general criticism of defining critical thinking as being comprised of a set of skills is that critical thought also requires particular dispositions or habits to use those skills. Dispositions, unlike skills, cannot be taught; they can only be cultivated through such activities as modeling. Many proponents of critical thinking classify both abilities and dispositions that are necessary in critical thought.

In light of these criticisms, Ennis revised his taxonomy to incorporate dispositions as well as specific abilities, thus defining critical thinking as a combination of the two (“Taxonomy”). He drew his list of abilities from the field of logic and these abilities are often taught in “informal logic” courses. The abilities are comparatively easy to evaluate (Ennis, “Assumption-Finding”). In fact, Ennis authored the Cornell Critical Thinking Test, which is a popular evaluative tool that measures critical thinking according to his taxonomy of abilities (Royalty). Dispositions are more difficult to measure.

Similar to Ennis, Paul created a scheme that addresses both abilities and dispositions, but Paul tended to stress the activities of the thinker more than the thought itself. He defined critical thinking as a “unique kind of purposeful thinking in which the thinker systematically and habitually imposes criteria and intellectual standards upon the thinking; taking charge of the construction of thinking; guiding the construction of thinking according to the standards; and assessing the effectiveness of the thinking according to the purpose, the criteria, and the standards” (Paul 21). In this way, Paul accounts for the dispositions of the thinker as well as requiring that the thinker master a certain skill set in identifying and using standards and criteria. His standards and criteria seem, like Ennis, to be primarily drawn from the field of logic.

At the far end of the spectrum are thinkers who identify critical thinking with the virtue of thoughtfulness (Schrag). The cultivation of thoughtfulness leads individuals to engage in purposeful and deliberate thinking (Schrag 7). Schrag defines a virtue as mediating between two contrary inclinations, in this case, between the inclination towards impulsiveness and the inclination towards rigidity (Schrag 14). If we ascribe to
Schrag’s notion that critical thinking is a virtue, then it is highly transferable. However, it is also exceedingly difficult to evaluate whether or not someone possesses a virtue. Although not explicitly stated, Schrag seems still to hold that deliberate thinking should be in alignment with the laws of logic, although his emphasis is on character building rather than skill-mastery. However, an important distinction is that, unlike skills, virtues cannot be taught; they can only be fostered. To accept Schrag’s proposition would entail that more attention be given to creating environments and situations that cultivate virtues rather than teaching them. In this he seems to be influenced by John Dewey’s ideas of reflective thinking (Dewey).

Critical vs. Creative Thinking

A significant criticism of all of these theories about critical thinking is that, drawing on the laws of logic, they focus almost all attention on teaching students how to evaluate propositions. This approach may help individuals decide what to do or believe, but it does not address what Nickerson calls “thinking at goals” or “thinking about value systems.” Nickerson is referring to an even higher level of critical thinking, not merely evaluating propositions, but thinking about what one’s goals should be. Such critical thinking also has a creative aspect that helps generate new propositions.

For all of the above thinkers, creative thinking is not in conflict with critical thinking, even if it is not part of it. Using Dewey, Schrag argues that thinking is an action that creates experience. Therefore, any thinking is creative and adds to the individual’s repertoire of experience that will generate new growth.

Some scholars, however, believe that critical thinking and creative thinking are different cognitive skills (Belenky; Walters). Part of their concern is that the emphasis on critical thinking privileges certain epistemologies at the cost of others. Walters places special value on imagination and intuition but is silent on the subject of whether or not they can be taught.

This position is most clearly articulated by the gleeful maverick Edward DeBono. DeBono’s main thesis is that so-called “vertical thinking” cannot construct new hypotheses but can only evaluate propositions. He creates a dichotomy between “vertical thinking” and “lateral thinking,” which might be called “creative thinking.” Lateral thinking involves creating whole new ways of looking at things, instead of evaluating old ways of looking at things. Additionally, DeBono claims that lateral thinking can be taught.

The creative/critical thinking dichotomy is problematic because new, creative insight is clearly essential to good thinking, yet how it occurs remains shrouded in mystery. Furthermore, creative thinking is clearly not tied to specific skill sets, nor is it readily identified by the possession of certain dispositions. Although creativity may be recognized, it is difficult to evaluate.

Is Critical Thinking Subject-Dependent?

In addition to the critique of the critical thinking movement privileging rational epistemologies at the expense of creative thinking, another essential debate within the movement is whether general critical thinking skills that are not subject dependent even
exist. McPeck, for example, argues that since thinking is always thinking about something, then any kind of thinking is dependent upon the subject being thought about. Furthermore, he claims that different subjects or domains have different epistemologies, the idea of critical thinking having different meanings from subject to subject. “[S]kills, like critical thinking in general, are parasitic upon detailed knowledge of, and experience in, parent fields and problem areas” (McPeck 10). He redefines critical thinking to mean reflective epistemology that is “the analysis of good reasons for belief, understanding the various kinds of reason involves understanding complex meanings of field-dependent concepts and evidence” (McPeck 24).

The ramifications of the proposition for subject dependence are enormous. If true, it undermines the claim that there are universal intellectual standards of reasonable thought. Moreover, it means that critical thinking can only be taught from within a particular subject, invalidating the many “critical thinking courses” that currently exist. Finally, the claim of subject specificity utterly negates all claims to the general transferability of critical thinking skills.

Paul and Ennis (“Critical Thinking”) criticize McPeck for reifying the concept of “subject,” which he uses interchangeably with the academic disciplines. Although one cannot think about “everything in general” but must think about a topic, that is not the same as thinking within a discipline, they argue.

Four possible overarching schemes for teaching critical thinking emerge from this debate. The “general” approach involves teaching generalized critical thinking skills in a critical thinking course. The “infusion” approach requires self-consciously teaching critical thinking skills from within a subject course. The “immersion” method assumes that students will gain the subject-specific critical thinking skills through taking the subject course. The “mixed-model” approach combines a general course with either an “infusion” or “immersion” approach (Ennis, “Critical Thinking”). Neither the “mixed-model” nor the “infusion” approaches rule out the possibility of generalizable critical thinking skills. Wilen emphasizes the role of metacognition in critical thinking and advocates an infusion approach in which the expert models critical thinking in light of their subject matter through thinking aloud. He endorses this teaching method not because critical thinking is restricted to subject matter but because such teaching is efficacious. Only through evaluating and comparing the general and the infusion methods will it be made apparent whether critical thinking skills are generalizable or subject-dependent.

**Evaluation of Critical Thinking**

Careful evaluation of critical thinking is vital, not only to answer the question of the generalizability, but also to assess program strengths and weaknesses and to redress the weaknesses. Baron identifies four dimensions in the evaluation of programs teaching critical thinking skills. The first axis is Formative-Summative. The Formative evaluation has to do with improving the program, while the Summative evaluates the effectiveness of the program. The second is Product-Process. The Product evaluation focuses on what the students produce while the Process is concerned with the workings of the critical
thinking instruction and thinking of the students. A third axis is QualitativeQuantitative. Both Qualitative and Quantitative studies are meant to capture the experiences of the people in the program. Finally, Experimental-Quasi-Experimental is another way of testing program effectiveness. Baron also suggests testing for sustained effect, transfer, side effects and metacognition.

A number of tests measure critical thinking skills, (Baron; Ennis, “Assumption-Finding”), but measuring dispositions and “virtues” is more difficult. A danger is that the ideals of critical thinking dispositions and Schrag’s idea of the virtue of thoughtfulness (Schrag) could be overlooked in application because evaluating skills is easier.

Halpern highlights a number of challenges in evaluating the effectiveness of critical thinking instruction. Pre-tests and post-tests convey some information, but reveal nothing about retention. Also, cognitive skills improve with practice, so presumably the real effects of critical thinking instruction would take some time to become apparent. Long-term gains are difficult to assess by testing students long after they have taken a critical thinking course, because it is difficult to measure whether any improvement from their baseline resulted from the course or from normal maturation and skills gained in other course. This problem of assessing program effectiveness is further complicated by the fact that virtually all campuses that offer critical thinking courses require them for all students; there is no control group. One possibility is to create a control group from students at a similar university.

There are also significant criticisms about the critical thinking instruments in addition to the fact that they only test skills and neglect habits, creativity and the “virtue of thoughtfulness.” McPeck argues that they are indistinguishable from IQ tests. One (Royalty) has been to test students using both IQ tests as well as critical thinking instruments. However, even if an immediate post-test does reveal information and is distinguishable from an IQ test, the program cannot be deemed successful until it can demonstrate transferability of the skills by the student into a new context.

**Transferability of Critical Thinking**

As previously mentioned, if critical thinking is subject-dependent, critical thinking should not be transferable across domains. If critical thinking is not transferable, then most of the reasons for teaching it are invalidated. However, other impediments may prevent the transfer of critical thinking skills. Perkins enumerates three stages of critical thinking development: acquisition, making it automatic, and transfer. Most critical thinking programs focus on acquisition, but without the other two steps, Perkins argues, critical thinking tends to remain within the context of the course. Perkins postulates two types of transfer. “High-road transfer” is the intentional transfer of a “frame” or critical thinking tactic from one learning context to another context. Instructors should create exercises that help students achieve this type of transfer because it typically does not occur automatically. “Low-road transfer” occurs more spontaneously. By “low-road transfer” he means the phenomenon of perceiving similarity in a new circumstance and applying the “frame.” The primary significance of Perkin’s thesis is that without instruction in transfer, students will be less likely to be able to apply critical thinking
skills to novel situations.

Some Studies of Critical Thinking Program Effectiveness and Transfer

Although most works dealing with critical thinking tend to be abstract and philosophical, there have been some notable attempts to evaluate the effectiveness of critical thinking instruction. One of the most impressive was conducted in Venezuela (Hernstein, et.al.). Four hundred seventh graders took a yearlong critical thinking course in which they were taught fifty-six lessons. Four tests were given before, during, and after the course. The same tests were administered to a control group. These tests were supplemented with a design and an oral argument post-test of randomly selected individuals from the experimental and control group. The results of one objective test were marginal, but the others were significant. The higher performance on the General Abilities Test and a Target Abilities Test developed for the program was especially notable. The two post-tests also revealed higher achievement in the experimental group. The tests measured critical thinking as skills. This study seems to strongly indicate transferability. The objective tests were cross-discipline and the oral arguments and design exercises cross domains.

Royalty created a study in an attempt to prove the generalizability thesis of critical thinking, and therefore, its transfer. He acknowledges that the lack of agreement about what constitutes critical thinking creates problems in its measurement. He conducted two studies to see whether or not critical thinking skills could be applied to novel domains. To measure this, he attempted to identify areas that did not require specialized subject knowledge in order to test them.

Royalty’s first study used the Cornell Critical Thinking Test, which was created by Ennis and reflects his theory of critical thinking. He also administered an IQ test, the Belief in the Paranormal Scale, and The Paranormal Experience Scale. His results showed a strong correlation between IQ and critical thinking but that neither critical thinking nor IQ accounted for the variance in belief in the paranormal. There was, however, a strong correlation between The Paranormal Experience Scale and the Belief in the Paranormal Scale. He explained his conclusion by postulating that belief in the paranormal and metaphysical speculation may rely upon other “ways of knowing” that are not part of critical thinking.

I believe that his first study was flawed because of some of the researcher’s assumptions. “Although the relationship between paranormal beliefs and experiences may represent logical consistency, it would represent critical thinking only if one discounted the importance of the content truth of the premise” (Royalty). Clearly, the researcher has already identified belief in the paranormal as illogical. He seems to be subscribing to a materialist ontology and a logical positivist epistemology. If Dewey is correct, then it is no surprise that experience of the paranormal is positively correlated with belief in the paranormal. It seems to me that his first study was flawed in its conception.

Royalty’s second study, however, is more promising. He administered the Cornell Critical Thinking Test, an IQ test, and a Statistical Reasoning test. Both IQ and
the critical thinking accounted for variability, but critical thinking accounted for a unique portion of the variability. Therefore, the test indicates that critical thinking skills can be applied to novel subject areas and, therefore, transfer.

In an attempt to understand transfer, Lehman and Nisbett examined three types of reasoning: verbal, statistical-methodological, and conditioned reasoning. They tested University of Michigan undergraduates during their first term in the first year, and then tested the same undergraduates during their second term in their fourth year. This study was looking primarily at the immersion model, not testing students who had formal critical thinking instruction. The students were given three tests, one for each reasoning type, that included both academic and “real life” questions. The first year showed no distinction based upon discipline major. After four years, there was no statistical significance shown for verbal reasoning. Those majoring in the Social Sciences or Psychology made significant gains in the statistical-methodological reasoning, while those majoring in the Natural Sciences or Humanities made marginal gains. The Natural Science and Humanities majors made significant gains in conditional reasoning, but the Social Science and Psychology majors did not improve. This study seems to support the conclusion that different types of reasoning are taught in different fields. Still, it does not rule out the possibility of general critical thinking skills. The improvement on the “real life” sections of the tests indicates that the skills were transferable, although different, depending upon the major in which the student was immersed.

In opposition to these studies, Hendricks studied critical thinking instruction looking at the distinction between traditional schooling that decontextualizes knowledge vs. situated learning. Two hundred and twenty seventh graders were studied, assigned randomly to experimental or control. The students were then taught about causality, an important component of critical thinking. They were given a “transfer task” to complete two weeks after the instruction and another six weeks after the instruction. Interviews were conducted after six weeks. Transfer was very poor for both groups of students. Hendricks suggests, in agreement with Perkins, that more direct transfer training was probably needed.

Conclusions and Needs for Additional Research

At this point, the evidence is mixed concerning the transferability of critical thinking skills and, therefore, the usefulness of critical thinking instruction. The confounded state of research into the transferability of critical thinking skills stems from fundamental disagreement about what is meant by critical thinking. All of the evaluation instruments test critical thinking as skills and do not test for dispositions nor for thoughtfulness. They are unable to account neither for “wrong” answers that might be reached through critical thought nor for “right” answers that might be reached through “test wiseness.”

The evidence also demonstrates the limitations of philosophy as a tool to explore these problems. Although philosophy is important to articulate goals and aims, all of the models have logical consistency and cannot be discredited using philosophical methods. I am most persuaded by the philosophy that defines critical thinking as requiring both
skills and dispositions. I do, also, acknowledge that different fields privilege different epistemologies, but I do not think that this discredits the existence of general critical thinking skills and am wary about reifying subject domains.

Clearly additional research is needed. Generally emphasis should shift from the philosophy of critical thinking to the evaluation of critical thinking and critical thinking programs. Detailed comparison between IQ test results and the Critical Thinking instruments is needed in order to ascertain if they are actually measuring different phenomena. Ways of studying critical thinking dispositions or thoughtfulness need to be created to ascertain whether or not these characteristics have been inculcated through critical thinking instruction. More quantitative and qualitative studies are needed to amass sufficient data for macroanalysis. The infusion model and the general model of critical thinking instruction should be subjected to tests and compared in order to ascertain whether general critical thinking skills exist or whether they are subject dependent. The impact of “transfer instruction” needs to be tested through more studies of transferability, including longitudinal studies and studies that measure the application of critical thinking to novel situations and new domains. Furthermore, the interaction between creative and critical thinking requires more exploration, including whether creative thinking, imagination and intuition can be taught. Of course, formative evaluative work to gauge the effectiveness of various methods of critical thinking instruction is still important.

**Works Cited**


