

Evaluation of Halpern's "Structural Component" for Improving Critical Thinking

Ana M^a Nieto and Carlos Saiz

Universidad de Salamanca

Halpern (1998) proposed a four-component model for promoting the transfer of critical thinking. One of them, the "structural component," focuses on how to organize teaching so that critical thinking skills can be generalized. Here, we assess the efficiency of that type of organization. Thus, one group of university students received instruction following the suggestions specified in that component and their performance was compared with that of other university students who received instruction in the same skills but using a different procedure, and with that of a control group. In comparison with the control group, the performance of both instructed groups was better after training. However, no significant differences were observed between either instruction group; both forms of instruction afforded very similar results.

Keywords: critical thinking, instruction, reasoning, decision-making

Halpern (1998) propuso un modelo de cuatro componentes para promocionar la transferencia del pensamiento crítico. Uno de ellos, el "componente estructural", se centra en cómo organizar la enseñanza para que las habilidades del pensamiento crítico puedan generalizarse. En este trabajo, evaluamos la eficiencia de este tipo de organización. Así, un grupo de estudiantes universitarios recibieron instrucción siguiendo las sugerencias especificadas en ese componente y su ejecución se comparó con la de otro grupo de estudiantes universitarios que recibieron instrucción en las mismas habilidades pero usando un procedimiento diferente, y con la de un grupo control. En comparación con el grupo control, la ejecución de ambos grupos instruidos fue mejor después del entrenamiento. Sin embargo, no se observaron diferencias significativas entre los dos grupos de instrucción; ambas formas de instrucción produjeron resultados muy similares.

Palabras clave: pensamiento crítico, instrucción, razonamiento, toma de decisiones

Correspondence should be addressed to Ana M^a Nieto, Departamento de Psicología Básica, Psicobiología y Metodología de las Ciencias del Comportamiento, Universidad de Salamanca, 37071 Salamanca, (Spain). Phone: 923- 294610- ext 3278. E-mail: acarracedo@usal.es

Improving critical thinking is undoubtedly one of the priority objectives of education at all levels. To train young people to be able to analyze, assess, judge, and construct information critically is just one of the challenges faced by educators in the 21st century. The broad consensus regarding the importance of critical thinking and the desire to improve it contrasts with the lack of critical thought about its conceptualization and how it should be taught.

In this sense, despite innumerable efforts that have been made to establish a clear and precise definition of critical thinking, none has been widely accepted. Some of the best known efforts are, for example, the proposal by McPeck (1981), who defined critical thinking as "The skill and propensity to engage in an activity with reflective skepticism" (p. 7) or that of Ennis (1987), who defines it as "reasonable, reflective thinking that it focused on deciding what to believe or do" (p.10), whereas for Lipman (2003), critical thinking is thinking that facilitates the undertaking of good judgments within and by criteria; it is self-correcting and sensitive to context. The nature of critical thinking is indeed so complex that it is not easy to synthesize it in a single definition, although such definitions do allow us to sketch some of its most important characteristics: skills, dispositions, reasoning, self-correction, judgments, decisions, beliefs, etc.

There is also much debate about whether the skills of critical thinking are general or, in contrast, they are specific and content-dependent. McPeck (1981, 1990) was one of the first to question the concept of critical thinking as a general skill, independent of content. He maintains that there are no generalizable thinking skills and considers that thinking is always "thinking about something" so that it is not possible to teach "thinking in general"; curricular disciplines are the best way to improve critical thinking and hence it should be taught within the various academic disciplines. This stance has been criticized by other authors (e.g., Govier, 1987; Paul, 1990; Siegel, 1990) who adopt the opposite position; that is, in defense of some general principles of critical thinking that can be taught independently of disciplines or specific areas of knowledge.

The main reason why people have gone against general courses on critical thinking (Brown, 1997; Glaser, 1984; Kurfiss, 1988; McPeck, 1981, 1990; Perkins, 1987) is due to the difficulty in achieving the transfer of skills learned in other domains, contents, or situations. Nevertheless, this is not a problem exclusive to general courses but also for the critical thinking disciplinary approach, as there are no guarantees that the skills of this, taught in a given discipline—say, History—will be applicable to another discipline—for example, a Science subject. In fact, one of the major difficulties inherent to programs designed to teach critical thinking is how to organize the teaching so that what has been learnt can be applied beyond the learning situation, in other contexts or in other situations.

It is widely agreed in the literature (Kennedy, Fisher, & Ennis, 1991) that the transfer of critical thinking from one domain to another is desirable and that approaches to teaching critical thinking should be designed to foster such transfer. In this sense, Halpern (1998) proposed a model for promoting the transcontextual learning of the skills of critical thinking so that the probability that these will be applied in real life, outside the class situation in which they were learned, is increased. The model is made up of four components. In the first two, Halpern specifies *what to teach* to improve critical thinking. In the other two, she offers a procedure for instruction; that is, *how* to organize the teaching so that what is taught is really learned and applied when the situation in hand so requires. Let us take a brief look at each of these components.

1. Instruction in and practice with the critical thinking skills. Halpern proposes the following category of skills for guiding instruction: (a) verbal reasoning, (b) analysis of arguments, (c) confirmation of hypothesis, (d) probability and uncertainty, and (e) decision making and problem solving.
2. Disposition to engage in a difficult type of thinking and its learning. It is important to distinguish between the capacity for thinking critically and the disposition to apply such skills. Some people may have great skills and yet lack the disposition to apply them and, hence, the dispositional component is very important. Among the dispositions pointed out by Halpern are the following: (a) the disposition to become committed and persist in a complex task, (b) a common tendency to forge plans and suppress impulsive activity, (c) flexibility and impartiality, and (d) the disposition to abandon unproductive strategies, etc.
3. The structural component to promote transfer. The final aim of teaching critical thinking is not only that students be able to understand and use the skills or strategies taught but also that they become able to use them in new situations when necessary. That is, there should be transfer or generalization of what has been learned in the class context to other situations. Halpern considers the problem of transfer as one of memory, as what is recalled depends on what happens in the learning process. Accordingly, Halpern suggests that learning should be organized so that it will not depend on content in order to facilitate the recovery of skills. With this in mind, she offers the following suggestions: (a) make the structural aspects of problems and arguments as relevant as possible so that learning them will not depend on the content; (b) encourage practice with different classes of examples; have the students do exercises and tasks similar to those found in the real world; (c) offer corrective feedback to develop the habit of "becoming aware spontaneously;" and (d) promote effective elaboration of information in memory, for example, by the use of organization so that interconnected

knowledge structures can be developed (the use of reflexive questions is a technique that can help achieve this) that will be similar to those found in daily contexts.

4. The last of the four components is meta-cognition, in which it is suggested that students should be faced with a series of issues to help them to convert implicit cognitive processes into explicit ones.

As stated, Halpern proposes this model for the teaching of critical thinking to promote its transfer across the different domains. Although the final goal of all instruction programs is transfer, its assessment (Perkins & Grotzer, 1997; Saiz, 2002) requires certain prior steps. Thus, before assessing the degree of transfer of an instruction program, it will be necessary to demonstrate that it produces some kind of improvement in the skill taught; that is, we must measure *the magnitude of the change*. Second, we must see whether the change persists over time, what is called (Perkins & Grotzer, 1997) *the degree of permanence*. Finally, it is necessary to assess whether the change observed is also produced in situations other than the teaching circumstance; that is, whether there has been *transfer*.

In the present work, our main aims are, first, to assess the efficiency of a general program for the improvement of critical thinking and, second, to evaluate the magnitude of the change produced by the "structural component" to foster the transfer, proposed by Halpern (1998). To achieve these aims, we developed a program for improving critical thinking in which we taught the following skills: (a) analysis and evaluation of daily arguments; (b) deduction and induction; (c) probability and uncertainty, and (d) decision making. Within the procedure for improving these skills we included the suggestions set forth in the structural component to promote transfer, proposed by Halpern and described above. This program was applied to a group of university students (the *structural group*), whose performance was compared with that of another group of university students to whom the same skills were taught but with a procedure that did not include the suggestions of the structural component (the *nonstructural group*), and with that of a control group who did not receive any type of instruction in critical thinking.

We formulated the following hypotheses: First, if the students are instructed in critical thinking (the structural and nonstructural groups), at the end of the treatment period, their performance should be better than that of the group that did not receive any kind of instruction (*control group*). Our second hypothesis was that if the students were instructed in critical thinking by means of the structural component approach (the structural group), which improves the conditions for learning critical thinking, at the end of the instruction period, they should show a better performance than those who received the instruction without this procedure (the nonstructural group).

Method

Participants

The definitive sample comprised 112 students with a mean age of 21.8 years from the third and fourth years of a Psychology degree (104 women and 8 men). The fourth-year students participated in one of the two instruction conditions, while those from the third year served as controls and received only the pre- and posttests, with no type of instruction in thinking during the period between both tests (8 months). Choice of the fourth-year students as the treatment group was based on academic reasons because in that year, the students take courses whose practical activities we could use for carrying out the instruction. Within this group, the students were randomly designated to one of the two experimental conditions. Despite the difference of one academic year between the control group and the two experimental groups, prior analysis of the means of the dependent variables allowed us to observe that there were no significant differences among them, so that starting level was the same for all of them. Accordingly, the groups could be considered perfectly comparable.

The sample in each group comprised 42 students who received the structural instruction, 36 who received training without that structural procedure, and 34 who served as controls.

Treatment

In the two groups of instruction, the same skills were taught. Generically, they were as follows: skills in analysis and evaluation of daily arguments; deductive (categorical and propositional) skills and inductive (causal, analogical and generalizations) skills, probability and uncertainty, and decision making (definition of the framework problem of the decision, the search for and evaluation of information, establishing conclusions and improvement of feedback). Both groups received the teaching of critical thinking in the same number of sessions (23 sessions of one hour weekly), with the same schedule and the same instructor. All students attended at least 90% of the sessions and had to hand in the tasks required; apparently this did not pose any problems for them. The instructor was the first author of this work.

However, the instruction procedure was different for each group and this is what really determined each of the experimental conditions.

Structural group: The instructors explained the skill being taught and, to do so, they used a greater number of examples of arguments, and strong emphasis was placed on the *structural aspects* of such arguments so that learning this skill would not depend on the content but on the structure. Following this, the students did exercises and tasks. In order to increase *the amount of practice* they were doing, a greater number of exercises were included than in

the other (nonstructural) group; also, these exercises were more realistic and similar to those found in the real world. The tasks performed by the students were collected and corrected by the instructor. This allowed us to analyze and assess the degree of understanding and learning achieved by the students. All participants received an analysis and assessment of their own work. Correction was completed with an overall feedback to the whole class in which the instructor explained all the doubts and mistakes in understanding, and the difficulties observed were clarified. In other words, the students received a corrective feedback about their learning and the knowledge they had gained. This technique is very similar to one of the Classroom Assessment Techniques (CATs) remarked by Angelo (1995) known as "Minute paper," which this author suggests could be used to improve meta-cognition. We used it with two goals in mind: first, that the students should have feedback about their performance and, second, that this would allow them to improve their meta-cognitive knowledge. In synthesis, this group received instruction following the suggestions gathered within Halpern's (1998) structural component as a way to improve learning.

Nonstructural group. In this group, the instructors explained the skill to be learned. They simply explained what the students had to do, placing less emphasis on the structure of the arguments and situations typically related to that skill. In other words, no reference was made to the structural characteristics of the arguments. The number of examples offered was lower than in the structural group. Following this, the students did exercises and tasks, whose number was also smaller than in the structural group and they were more "formal" or "academic" and less related to daily situations. The students performed these exercises through an interactive program developed by us (<http://web.usal.es/~emid/pips/pips.htm>—the access keys for each block can be consulted in Appendix A). As they were performing the exercises, the computers gave them feedback about their performance, in the sense of whether the answers were correct or incorrect, but no explanation was given as to what they had done wrong or why it was incorrect. Thus, they were offered immediate feedback but this was not corrective. Accordingly, the students from this group were working the skills with no emphasis on structural aspects, with few examples, with less practice, with fewer common daily situations, and with noncorrective feedback.

Control group. The students in this group did not receive any kind of instruction in critical thinking either within or outside their curricular activities, and they only performed the pretest and the posttest.

Evaluation Instruments and Correction Criteria

To determine whether the instruction improved critical thinking, we decided to apply several tests. Thus, we chose a standardized test such as the Cornell-Level Z Test (Ennis

& Millman, 1985), which provides information about the main critical thinking skills such as deduction, induction, credibility, etc. However, this test does not assess the decision-making skill, so we decided to elaborate an instrument for this. Another very important skill in critical thinking, which is not evaluated directly in the Cornell Test either, is that of analyzing and assessing arguments, essentially informal ones. Therefore, we developed a Reasoning Test so that this aspect would also be covered. Finally, we applied the Knowledge Test as a complementary one; this would provide us with information about the beliefs about and attitudes towards thinking. They are described in some detail below.

In the pretest, the following assessment instruments were used:

The Cornell Critical Thinking Test – level Z, (Ennis & Millman, 1985), which comprises 52 items with three alternative answers to each. It evaluates the following skills: induction, deduction, observation, credibility, assumptions, and meaning. The main purpose of this test is to obtain a standardized measure of critical thinking. The test has a reliability of .78.

Knowledge Test. This comprises 80 questions with true-false answers. The items were the issues raised by Halpern (2003) in the different chapters of her manual on critical thinking. The aim of this test was to obtain a measure of students' beliefs about certain aspects and situations related to thinking. Thus, the test evaluates certain ideas about thinking skills, the use of certain strategies to think better, and the perception or lack of perception of certain biases or errors, etc. The reliability index is .65

Reasoning Test. We developed this instrument to assess the skills of identification, analysis, and evaluation of arguments. It comprises 10 texts, among which there are arguments, sound or unsound, and nonarguments. For each of the texts, the students were asked to state whether the text was an argument; if it was, they had to identify the conclusion and the reason(s) offered, and they had to make an assessment of the argument, in the sense of whether they agreed with it and why. The correction of this test was carried out based on the following criteria: one point was awarded if arguments were correctly differentiated from nonarguments; if the conclusion was correctly identified, two points were awarded; and for each reason correctly identified, another point was awarded. For each additional correctly assigned element (assumptions or counterarguments), another point was awarded. If the students perceived relationships among the reasons, one, two, or three points were awarded, depending on the difficulty. Finally, if the students evaluated the argument correctly, two points were awarded. The reliability index is .65.

Decision-Making Test. We also developed this instrument to evaluate the main skills involved in decision-making. It was made up of 21 problems, of which 20 aimed at evaluating Vulnerability to bias, and one

evaluated the actual decision process. In the 20 problems aimed at evaluating Vulnerability to bias, we posed questions that would reflect the most typical errors committed in decision-making. In particular, we evaluated the following aspects: the construction of scenarios; the framework of the decision; the heuristics of availability, of representativeness, of anchoring and adjustment; overconfidence, anchoring, and confirmatory bias. With the exception of the questions exploring overconfidence and regression to the mean, all were multiple-choice. The correction criterion used for these questions was as follows: if the students committed biases, they were awarded 0 points and, if not, 1 point. The last question, which evaluated the Decision-making process, was open. The students were asked to explain how they would make a decision. The aim was that they should at least specify the main phases of the process (reflecting on what is desired and defining the problem, searching for information, generating alternatives, choosing one, and assessing the choice and process used). Ultimately, this question was not included in the final analyses, as it had very low reliability indexes and negative correlations with some of the measures. The reliability of the test, excluding this question, is .42.

The correlations between these tests before instruction ranged between .096 ($p = .316$) and .291 ($p < .05$). All correlations were significant, with the exception of the former one, corresponding to the correlation between the Reasoning Test and the Cornell Test which, as seen, was nonsignificant.

At posttest, we applied the same tests: Cornell Test (Cronbach's $\alpha = .73$), Knowledge Test (Cronbach's $\alpha = .64$); for the Reasoning Test and the Decision-making Test, parallel forms were applied, and the respective reliability indexes were .63 and .61.

The correlations between the posttests were between .165 and .471. All were significant, with the exception (as at pretest) of that obtained between the Reasoning Test and the Cornell Test, which was nonsignificant ($r = .165$, $p = .08$).

The correlations between pretest and posttest measures were: .426 ($p < .01$) for the Cornell Test, .401 ($p < .01$) for the Knowledge Test, .294 ($p < .01$) for Vulnerability to Bias Test, and .155 ($p = .103$) for the Reasoning Test.

Procedure

One week before starting the instruction, we applied the pretest exercises to all the students, and one week after it had been completed, the posttest was implemented. For both instruction groups—structural and nonstructural—critical thinking was taught over 23 weeks, at one hour per week. In both groups, the time dedicated to each of the skills was exactly the same, the only difference being the way of teaching.

Results

Table 1 shows the means and standard deviations of the three groups: structural, nonstructural, and control, for each of the dependent variables before and after instruction. As the control group included students from one academic year lower than those in the two groups receiving instruction, we decided to perform a single-factor ANOVA to verify that there were no differences among the three groups before treatment. The results showed that all three groups had a similar result in critical thinking as evaluated with the Cornell Test, $F(2, 109) = 0.45$, $p = .63$; their knowledge of critical thinking, as evaluated by the Knowledge Test, $F(2, 109) = 0.157$, $p = .85$; in the number of biases committed, assessed by the Vulnerability to Bias Test, $F(2, 109) = 1.68$, $p = .19$; and in the Reasoning Test, $F(2, 109) = 2.42$, $p = .094$. Once we had verified that all the students had a similar level in these critical thinking skills, we evaluated the quality of the instruction.

To evaluate the effectiveness of the two methods for teaching critical thinking in comparison with the control group, we performed a multiple analysis of covariance,

Table 1
Pre-Posttest Means and Standard Deviations

GROUPS	Cornell Test	
	Pretest <i>M</i> (<i>SD</i>)	Posttest <i>M</i> (<i>SD</i>)
Control	15.29 (6.18)	16.64 (5.13)
Structural	17.21 (6.67)	19.41 (6.57)
Nonstructural	16.18 (5.95)	19.63 (6.87)
Knowledge Test		
	Knowledge Test	
	Pretest <i>M</i> (<i>SD</i>)	Posttest <i>M</i> (<i>SD</i>)
Control	25.91 (9.23)	26.06 (8.35)
Structural	26.38 (8.14)	34 (9.96)
Nonstructural	27.06 (8.52)	33.53 (9.72)
Reasoning Test		
	Reasoning Test	
	Pretest <i>M</i> (<i>SD</i>)	Posttest <i>M</i> (<i>SD</i>)
Control	3.77 (1.37)	2.42 (1.10)
Structural	4.22 (1.66)	6.95 (1.93)
Nonstructural	4.60 (1.62)	5.85 (1.75)
Vulnerability to Bias		
	Vulnerability to Bias	
	Pretest <i>M</i> (<i>SD</i>)	Posttest <i>M</i> (<i>SD</i>)
Control	6.35 (2.159)	8.21 (1.68)
Structural	7.1 (1.81)	11.69 (1.90)
Nonstructural	7.03 (1.69)	11.22 (2.05)

incorporating the scores from the pretest as covariates so that we would be able to control the effect of undesired factors that might affect this quasi-experimental design. The analyses carried out using the Levene test and the Box test confirmed that the variances were homogeneous.

The analyses confirmed that, after the instruction, the means of the three groups differed significantly, Wilks' $\lambda = 0.436$, $F(4, 102) = 32.96$, $p < .01$, $\eta^2 = 0.564$, Power = 1. With the univariate contrasts, we observed that the differences among them occurred in the four dependent variables: in the Cornell Test, $F(6, 105) = 5.50$; $p < .01$; $\eta^2 = 0.239$, Power = 0.99; in the Knowledge Test, $F(6, 105) = 8.98$, $p < .01$, $\eta^2 = 0.339$; Power = 1; in the Vulnerability to Bias Test, $F(6, 105) = 14.98$, $p < .01$, $\eta^2 = 0.460$, Power = 1; and in the Reasoning Test, $F(6, 105) = 24.06$, $p < .01$, $\eta^2 = 0.579$, Power = 1.

The univariate tests revealed that, in the Cornell test, significant differences were seen between the nonstructural group and the control group, but not between the structural group and the control group. Accordingly, the instruction based on Halpern's theoretical principles did not lead to more improvement of the students' results than in the group not receiving instruction; both displayed a similar performance. Nor were any difference observed between both types of instruction: structural and nonstructural.

However, in the Knowledge Test, we did observe statistically significant differences between both types of instruction and the control group, so that the students who had received instruction in critical thinking using both the structural and the nonstructural approaches, proved to have a better level in their knowledge of critical thinking than those who did not receive instruction. Nevertheless, and contrary to our predictions, the structural group did not show better results than the nonstructural group, that is, instruction based on Halpern's principles did not contribute to an improvement in knowledge about thinking. Regarding Vulnerability to Bias, we observed that the two groups that had received instruction in critical thinking had better results than the group that did not receive any type of instruction. In other words, they committed fewer biases and mistakes, although again it was observed that the group that received structural instruction did not show a greater decrease in the number of biases than those who received nonstructural instruction. Regarding this dependent variable, the instruction according to Halpern's theoretical model did not produce better learning either. Finally, in the context of the Reasoning Test, both instruction groups showed significant differences with respect to the group not receiving any type of intervention and, additionally, the students who received instruction according to Halpern's suggestions demonstrated better skills in analysis and the assessment of arguments than the group that received instruction not incorporating her theoretical suggestions.

Discussion

Overall, our first hypothesis was confirmed. That is, the two groups that received instruction in critical thinking obtained better results than the control group, with the exception that, in the Cornell Test, the structural group displayed the same learning as the control group. As regards the other dependent variables, the students receiving instruction performed better than the group that did not receive it. Accordingly, instruction in critical thinking, following either of the two teaching procedures, was effective, although with respect to the Cornell Test, the nonstructural instruction was more effective. However, the second hypothesis was only partially confirmed. Thus, the students who received the structural instruction scored higher than those who received nonstructural instruction only on the Reasoning Test, but not as regards the other dependent variables. Accordingly, we can confirm that, at least at short-term, instruction that emphasizes the structural aspects of arguments, with a greater number of exercises similar to those found in daily life, and that provides corrective feedback, does not produce greater improvements than teaching critical thinking without such characteristics.

The students subjected to instruction in critical thinking displayed better skills in this after the instruction, whereas the control group did not improve in this respect. This was not the case of the Cornell Test, in which the group that received instruction following Halpern's theoretical principles did not show improved learning compared with the control group. We do not have a clear explanation for this. In this sense, our results are also similar to those obtained in other attempts to improve thinking, such as those obtained with the Intelligence Project (Hernstein, Nickerson, Sánchez, & Swets, 1986) and those obtained with Productive Thinking (Treffinger, Speedie, & Bruner, 1974) as the gains achieved by the groups receiving instruction were not revealed in all the measures. Specifically, the greatest gains in the instructed groups occurred in the tests designed by the authors to evaluate the skills taught, whereas in the standardized tests, the gains were smaller. In other words, this seems to be very commonplace in attempts to improve thinking and leads us to another series of issues. Are such tests really sensitive to the changes that are assumed to occur when instruction in thinking is taught? It may be that our instruments were flawless as regards their psychometric characteristics but were unable to detect the change; that is, they were insensitive to it. In fact, some authors (Ennis, 2003; Govier, 1987) have reported that multiple-choice tests might not be the ideal way to evaluate thinking. Instruction produces a series of changes in the thinking process that may be difficult to detect with a multiple-choice format. It seems that the best way to evaluate thinking is through open questions, in which students have to justify their point of view (Ennis, 2003). Nevertheless, and for this reason they are less frequently used, such tests are much more costly to correct.

In fact, with a view to solving certain problems detected in traditional tests on critical thinking, Halpern (2006) has recently developed an instrument—the Halpern Critical Thinking Assessment Using Everyday Situations—that is currently being adapted to the Spanish population and that combines two response formats: open and closed. The test is promising, although some time will be required before we can see how it really works. In synthesis, what does seem clear is that much more empirical work is required to determine whether the partial improvements observed are due to the instruction procedures, which failed to improve some skills, or to a problem with the assessment instruments that were insensitive to the changes.

Another aspect to be addressed is the absence of significant differences between the two instruction groups, as differences were only seen between them on the Reasoning Test, but not in the other dependent variables. Does this mean that instruction with Halpern's method is not more effective? First, it should be recalled that the recommendations made by Halpern (1998)—highlighting the importance of doing a lot of practice, with examples and exercises similar to those found in real life, underlining the structural aspects of arguments so that their learning will not depend on the content, and offering appropriate feedback—are aimed at *promoting the transfer* of thinking skills to contexts other than those explored in the instruction. Thus, it is possible that the instruction given to the structural group, which received instruction according to this approach, might not elicit short-term improvements, although the improvements gained would be longer lasting and would become manifest more clearly in daily contexts. That is, there would be greater transfer. Nevertheless, this possibility, related to a potentially greater persistence and a greater transfer, must be evaluated empirically. The assessment of a program aimed at improving thinking (Perkins & Grotzer, 1997; Saiz, 2002) begins with the demonstration of its immediate effects, and indeed this was our main goal here. And only when the change has been demonstrated is it possible to evaluate the degree of persistence and, finally, the degree of generalization.

However, regarding our results, the fact that the structural group displayed better skills in analysis and the evaluation of arguments, assessed in the Reasoning test, is a small proof of an immediate effect although, as stated above, it would be necessary to verify whether such efficacy persists in the long term and in other situations.

Nevertheless, it is also necessary to consider the gains obtained by the nonstructural group, which received instruction in critical thinking without application of Halpern's principles, in a positive light. The change observed in this group is relevant because it is proof that it is indeed possible to improve this type of thinking with a computer program and without directed, supervised instruction. This evidently implies a huge saving in time and energy. In fact, one work (Halpern, 2006) revealed improvements in critical thinking after this was taught on the web. Nevertheless, as

we have seen, it would be necessary to go deeper into the practical significance of the change and the persistence and generalizability of its effects.

To conclude, in contrast to some reviews (Ten Dam & Volman, 2004; Tsui, 1999) in which the authors highlight the few positive results of instruction in critical thinking and question the possibility of improving it by means of general courses, we believe that our modest results are promising and offer some evidence, although for the short-term, of the possibility of improving critical thinking with this general procedure. Moreover, whereas there are more reliable evaluation instruments, we feel we should continue searching for new and better evaluation instruments in order to be able to evaluate the skills we consider important in critical thinking but which have not been included in any standardized test available to date.

References

- Angelo, T.A. (1995). Classroom assessment for critical thinking. In D.F. Halpern & S.G. Nummedal (Eds.), *Psychologists teach critical thinking* [Special Issue]. *Teaching of Psychology*, 22, 6-7.
- Brown, A.L. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist*, 52, 399-413.
- Ennis, R.H. (1987). A taxonomy of critical thinking dispositions and abilities. In J.B. Baron, & R.J. Sternberg (Eds.), *Teaching thinking skills* (pp. 9-26). New York: Freeman.
- Ennis, R.H. (2003). Critical thinking assessment. In D. Fasko, Jr. (Ed), *Critical thinking and reasoning: Current research, theory, and practice* (pp. 293-313). Cresskill, NJ: Hampton Press.
- Ennis, R.H., & Millman, J. (1985). *Cornell Critical Thinking Test, Level Z*. (3rd ed.). Pacific Grove, CA: Midwest.
- Ennis, R.H., Millman, J., & Tomko, T.N. (1985). *Cornell Critical Thinking Test, Level X & Level Z-Manual* (3rd ed.). Pacific Grove, CA: Midwest.
- Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist*, 39, 93-104.
- Govier, T. (1987). *Problems in argument analysis and evaluation*. Dordrecht, Holland: Foris.
- Halpern, D.F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist*, 53, 449-455.
- Halpern, D.F. (2003). *Thought and knowledge: An introduction to critical thinking* (4th ed.). Hillsdale, NJ: Erlbaum.
- Halpern, D.F. (2006). *Halpern Critical Thinking Assessment Using Everyday Situations: Background and scoring standards*. Unpublished report.
- Herrnstein, R.J., Nickerson, R.S., Sánchez, M., & Swets, J.A. (1986). Teaching thinking skills. *American Psychologist*, 41, 1279-1289.
- Kennedy, M. Fisher, M.B., & Ennis, R.H. (1991). Critical thinking: Literature review and needed research. In L. Idol, & B. Fly Jones (Eds.), *Educational values and cognitive instruction: Implications for reform* (pp. 11-40). Hillsdale, NJ: Erlbaum.

- Kurfiss, J.G. (1988). Critical thinking: Theory, research, practice, and possibilities. *ASHE-ERIC Higher Education Report Nr. 2*. Washington, DC: George Washington University.
- Lipman, M. (2003). *Thinking in education* (2nd ed.). Cambridge, MA: Cambridge University.
- McPeck, J.E. (1981). *Critical thinking and education*. Oxford, UK: Martin Robinson.
- McPeck, J.E. (1990). *Critical thinking: Dialogue and dialectic*. New York: Routledge.
- Nickerson, R.S., Perkins, D.N. & Smith, E.E. (1985). *The teaching of thinking*. Hillsdale, NJ: Erlbaum. (Spanish translation: L. Ruano & C. Ginard. *Enseñar a pensar: aspectos de la aptitud intelectual*. Madrid: Paidós, 1994).
- Paul, R. (1990). McPeck's mistakes. In J.E. McPeck (Ed.), *Critical thinking: Dialogue and dialectic* (pp. 102-111). New York: Routledge.
- Perkins, D.N. (1987). Thinking frames: An integrative perspective on teaching cognitive skills. In J.B. Baron, & R.J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 41-61). New York: Freeman.
- Perkins, D.N., & Grotzer, T.A. (1997). Teaching intelligence. *American Psychologist*, 52, 1125-1133.
- Saiz, C. (2002). Enseñar o aprender a pensar. *Escritos de Psicología*, 6, 53-72.
- Siegel, H. (1990). McPeck, informal logic, and the nature of critical thinking. In J.E. McPeck (Ed.), *Critical thinking: Dialogue and dialectic* (pp. 75-85). New York: Routledge.
- Ten Dam, G., & Volman, M. (2004). Critical thinking as a citizenship competence: Teaching strategies. *Learning and Instruction*, 14, 359-379.
- Treffinger, D.J., Speedie, S.M., & Bruner, W. (1974). Improving children's creative problem solving ability: The Purdue Creativity Project. *Journal of Creative Behavior*, 8, 20-30.
- Tsui, L. (1999). Courses and instruction affecting critical thinking. *Research in Higher Education*, 40, 185-200.

Received January, 17, 2007

Revision received October, 9, 2007

Accepted October, 22, 2007

APPENDIX A

Access to the simple critical thinking program:

Internet site: <http://web.usal.es/~emid/pips/pips.htm>

Once the site has been entered, click on “Subject” (*tema*) and, in “Subject,” select “Thinking” (*pensamiento*), in which the corresponding topics appear. Access to each one requires codes, as below:

- Practical or informal reasoning: p67e45; pu84dt (with solutions)
- Syllogistic reasoning: htbs749
- Propositional reasoning: sirjt86
- Wason task: widkn10
- Inductive reasoning: dkring78.

Likewise, it is possible to access the decision-making block, except that this time, instead of selecting “thinking,” you select “decision-making.” The access codes are as follows:

- Heuristics: 2000
- Probability: 1000
- Decision tables: 3456