Revista Electrónica de Metodología Aplicada
2009, Vol. 14 nº 1, pp. 1-15

Analysis of the psychometric properties of the Spanish version of the HCTAES- Halpern Critical Thinking Test Using Everyday Situations

Ana María Nieto, Carlos Saiz y Begoña Orgaz[1]

Facultad de Psicología. Universidad de Salamanca.

ABSTRACT

The present study analyses the psychometric properties of the Spanish version of Halpern's Critical Thinking Test Using Everyday Situations. For the assessment of this type of thinking, it uses familiar situations with a double response format: open and closed. The test was applied to 335 Spanish students: 238 University students from the 3rd and 4th years of the degree course and 52 high school students. We analysed the reliability of the test and performed an exploratory factor analysis to assess its construct validity. The results suggest that the Spanish version of the test is a trustworthy instrument for assessing critical thinking. However, its factor structure is not strongly related to the theoretical structure of the test. Our results, rather than weakening the construct validity of the test, reflect the strong interrelationship among the different skills of critical thinking.

Keywords: Assessment, critical thinking, reliability, factor analysis, validity.

[1] Adress:
Ana Mª Nieto.
Departamento de Psicología Básica, Psicobiología y Metodología de las CC del Comportamiento.
Universidad de Salamanca, 37071 Salamanca, España.
Tlno: 923294610- ext 3278; E-mail: acarracedo@usal.es

1.- Introduction

At the historical moment at which we find ourselves thanks to technological advances we now only need a few seconds to access huge amounts of information, which must be converted into knowledge. To accomplish this, people need cognitive skills such as how to discriminate relevant from fallacious information, elaborate and assess arguments, make correct judgments about probabilistic events, perceive and recognize covariations, analyze causal relationships, develop good decision-making strategies, have the resources for solving certain problems, etc. All these skills form part of what is known as critical thinking. But what is critical thinking? Critical thinking is reasoned and reflexive thought focused on deciding what to believe and what to do (Ennis, 1996). It is not an automatic type of thinking, no mechanical, but quite the opposite, since it is intentional, is based on reflection and it is propositive. It is directed towards specific objects such as deciding upon our beliefs and actions when faced with certain problems or situations. According to Halpern (2006) "…It is the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions. Critical thinkers use these skills appropriately, without prompting, and usually with conscious intent, in a variety of settings. That is, they are predisposed to think critically. When we think critically, we are evaluating the outcomes of our thought processes--how good a decision is or how well a problem is solved. Critical thinking also involves evaluating the thinking process…" (p.3). It is unanimously accepted that critical thinking is formed of a set of skills and of a set of dispositions. The skills represent the cognitive component and the dispositions the motivational one. This distinction is very important because it reflects the fact that if a person knows which skill to apply in a given situation but is not prepared to activate it, that person will not exhibit critical thinking. It is important for people both to know how to apply their skills and also to wish to do so. The way of categorizing the skills of critical thinking varies considerably among the different authors who have addressed the issue. For example, Ennis (1987) distinguishes among the following skills: focusing on the issue at hand, analyzing arguments, asking and answering clarifying and/or challenging questions, judging the credibility of sources, defining terms, identifying assumptions, deciding and interacting with others. Swartz and Perkins (1990) proposed much more general categories, such as creative thought, critical thinking, decision making, solving of daily problems, and solving of mathematical problems. As may be seen, there is therefore no consensus as regards which cognitive skills really do form critical thinking. This lack of agreement was explored by a panel of experts (APA, 1990) who sought to reach agreement about the concept and meaning of critical thinking. The panel of experts identified the following skills as central for critical thinking interpretation, analysis, evaluation, inference, explanation and self-regulation. Although this categorization seems to be exhaustive, it perhaps suffers from the problem of being less operative when teaching and assessing critical thinking.

Improvement of this thinking has and continues to be a constant source of worry in educational contexts. In Spain, the new educational reforms are beginning to underscore the importance of training people to think in a critical way. For example, some of the Transversal Competencies set forth in the European Higher Education Area (EHEA) (Delgado, et al., 2005) are the capacity for analysis and synthesis, information management skills, problem solving, decision-making, the ability to criticize, the ability to generate new ideas… all of them forming part of what can be understood as critical thinking (e.g., Halpern, 2003a).

Working to improve critical thinking and the development of such activity implicitly demands that good instruments be available for assessment that will allow us to check

whether the teaching has been effective. An array of tests have been designed to accomplish this (Ennis, 2003), its differ depending on the type of population in which they are implemented, in the type of skills they evaluate, in the format of the questions, in whether they assess critical thinking as a general skill or within some academic area, etc. Among those most used are the Cornell Critical Thinking Test, Level X and Level Z (Ennis and Millman, 1985), the California Critical Thinking Skills Test (Facione et al., 1990), the Watson-Glaser Critical Thinking Appraisal Test (Watson and Glaser, 1980) and the Ennis-Weir Critical Thinking Essay Test (Ennis and Weir, 1985). All these tests assess this type of thought independently of academic disciplines. In respect of the format of the questions, in the first three the multiple choice questions are closed, while in the last one the questions are open and the individual participating in the test must offer an argument, opinion, etc. Regarding psychometric properties, the manuals of the tests offer data concerning their reliability and validity that justify their use. Despite this, there are some studies (Jacobs, 1995; Loo & Torpe, 1999) that have cast doubt on them. Thus, critical thinking, which analyses, evaluates, judges, questions and reasons, is hard to assess by means of questions alluding to invented or somewhat unreal situations, in which –additionally- the individual is forced to chose from a set of alternatives with no possibility of making any contributions or engaging in constructive searches, reflective thought, seeking inferences, etc., beyond what is presented in the material. In this sense, Govier (1987) pointed out that critical thinking is not easy to induce in tests with multiple alternative answers in which this type of thinking is addressed through a series of simple, clearly stated topics, which in most cases tend to be contrived and artificial, expressed in short sentences and as neutrally as possible so that they will not be susceptible to different interpretations This means that there are many aspects of critical thinking that cannot be measured with this type of test format. In turn, the latter suggests that critical thinking should be evaluated using open, argumentative questions and in interviews. Evidently, the costs are much higher. Along the same line, Ennis (2003) suggests that possibly the best format for the assessment of critical thinking would be one in which the individual would have to chose from among a series of alternative items and explain why that particular one, and no other, was chosen.

Over the years, Halpern (2003b, 2006) has been working on an instruments designed to assess critical thinking; to a large extent, she has been trying to remedy some of the problems addressed above. Her work has culminated in the development of the HCTAES –the Halpern Critical Thinking Assessment Using Everyday Situations- This test has certain characteristics that make it completely different from others currently in use (e.g., Ennis and Millman, 1985; Facione et al., 1990). First, it uses everyday situations, similar to those found in real life, such that the materials are ecologically valid since they are representative of what might be found in a newspaper or in a day-to-day discussion. Second, it uses a double format of questions. Thus, a situation or problem is presented and the answer to an open question is requested, after which the individual is asked to choose the best alternative that will solve it. According to Halpern, this dual format of questions on the one hand allows us to know, for example, whether the person answering the question shows spontaneous use of the skill of critical thinking and, on the other, whether that person is capable of using it when he/she is told that it is necessary for a given situation, even though he/she has not spontaneously recognized that it is necessary.

As mention above, critical thinking is the general rubric of the different cognitive skills forming it. Halpern (1998, 2003a, 2006) advocates a critical thinking composed of 5 major skills: Hypothesis Testing Skills, Verbal Reasoning Skills, Argument Analysis Skills, Likelihood and Uncertainty Skills, and Decision-Making and Problem Solving Skills. Thus,

according to this theoretical model she has developed the HCTAES, which attempts to assess those 5 skills through 25 scenarios or situations; 5 for each of the skills. Each of the scenarios is evaluated by means of a question in open format and a closed question. Its trustworthiness and validity are based on a series of studies (Halpern, 2006) although this author underscores the pressing need to replicate them.

The HCTAES has been translated into Spanish and is now being adapted by us for the Spanish population. We have undertaken a study with several samples of University students and students from the second year of High School. Our aims are to analyze the psychometric properties of the Spanish version by means of item analysis, together with an analysis of the overall reliability of the test and its construct validity through an exploratory factor analysis. In particular, we wish to check whether the structure of the Spanish version of the test matches that of the theoretical model proposed by Halpern.

2.- Method

2.1.- Participants

A total of 335 individuals participated in the study: 52 High-School students, 145 Psychology students from the Catholic University in Salamanca (73 from the 3rd year and 72 from the 4th year), and 138 students from the 4th year of the Psychology degree of the University (Laic) of Salamanca. The mean age of the participants was 21.03 (Sx =2.89) and 84% were women and the rest men.

2.2.- Instruments

We used the HCTAES. We have already explained that comprises 5 skills of critical thinking: Hypothesis Testing Skills, Verbal Reasoning Skills, Argument Analysis Skills, Likelihood and Uncertainty Skills, and Decision-Making and Problem Solving Skills. Its presented the students with daily situations similar to those seen in real life and offered two possibilities for answering them: one of them was open, in which the respondent had to provide an argument, an explanation or generate the solutions to the problem, and the other one was closed, in which the participant had to choose the best one from among a series of alternatives. The test was translated following a translation/back-translation procedure, employed by experts in Critical Thinking and in English, both Spanish and Latin-American

The scoring criteria for the open part, generically, are as follows: when the participant offers the correct response he/she receives two points; when the answer approaches the right one but this is not stated explicitly (that is an answer is given but in a less “technical” way), he/she is awarded one point, and when the answer is wrong the participant receives no points. In the closed part, the participant is asked to choose the correct alternative(s) and receives one point for each correct answer.

Below we show one of the questions used in the test, followed by the criteria for its scoring offered in the test

(1) Open part

A recent report in a magazine for parents and teachers showed that adolescents who smoke cigarettes also tend to get low grades in school. As the number of cigarettes smoked each day increased, grade point averages decreased. One suggestion made in this report was that we could improve school achievement by preventing adolescents from smoking.

In light of this information, **would you support this idea as a means to improve the school grades of adolescents who smoke.**
Yes
No
Please explain why or why not

Once you have answered, please go to the next page

(1) Closed part

A recent report in a magazine for parents and teachers showed that adolescents who smoke cigarettes also tend to get low grades in school. As the number of cigarettes smoked each day increased, grade point averages decreased. One suggestion made in this report was that we could improve school achievement by preventing adolescents from smoking.
Based on this information, **what would be the best answer** (choose one)

a) School grades probably will improve if we prevent adolescents from smoking because the researchers found that when smoking increases, grades go down.

b) School grades might improve if we prevent adolescents from smoking, but we cannot be certain because we only know that grades go down when smoking increases, not what happens when smoking decreases.

c) There is no way to know if school grades will improve if we prevent adolescents from smoking because we only know that smoking and grades are related, not whether smoking causes grades to change.

d) There will probably be no effect on grades if we prevent adolescents from smoking because the magazine is written for parents and teachers, so it is probably biased against adolescent smoking.

Skill: Hypothesis Testing **(1) Open Part : hypothesis** Yes: 0 points; **No: 1 point; Why? 2 points-** Correlational variables but showing that the correlation does not involve causality. **1 point**: It is stated that a third factor could be important but not necessary **0 points**. Irrelevant answers. E.g., smoking is bad for your health etc.. Maximum 3 points	**(1) Closed part:** a.o b.o **c.1** d.o Max 1 point

2.3.- Procedure

The participants completed the HCTAES test on a written format and the maximum time allowed was 120 minutes; the session was split into two parts to prevent fatigue.

3.- Results

We analyzed the reliability of the test. We first estimated reliability by skills and, as may be seen in table 1, this ranged between 0.34 on the Likelihood and Uncertainty skill and 0.63 of the Hypothesis Testing skill. We then calculated the overall reliability of the test using the Cronbach alpha (0.774), observing that the test had a good reliability index.

Table 1 Cronbach alpha by dimensions.

Dimension	Nº of elements	Cronbach alpha
Hypothesis testing	10	.635
Verbal reasoning	10	.367
Arguments analysis	10	.465
Likelihood and uncertainty	10	.340
Decision making and problem solving	10	.487

To analyze construct validity and check whether the structure of the test fitted Halpern's theoretical model comprising the 5 above-mentioned skills, we performed an Exploratory Factor Analysis. Once we had obtained the pertinent indices of the Kaiser-Mayer-Olkin test, and the Sphericity Test of Bartlett ($\chi 2= 2214.44$, $p< 0.05$), we analyzed the Principal Components. We observed that there was no unidimensionality, since not all the items had weights greater than 0.30 on the first component. The lack of unidimensionality supports a model of critical thinking comprising a whole array of skills, which is the notion most widely held in the field (e.g. APA, 1990; Ennis, 1996; Halpern, 1998; 2003a; 2003b Swartz and Perkins, 1990). Second, 19 factors were obtained (eigenvalues greater than 1), which accounted for 58.49% of the variance. This structure departs from that proposed by Halpern, formed by 5 factors, to a considerable extent. Figure 1 shows the Scree-test, in which it is possible to note the factor structure obtained with the items. With a view to obtaining a factor structure similar to that proposed by Halpern, we decided to perform a refactoring to five components on those 19 initially obtained factors. However, upon observing the last factors of the 19 obtained initially, we found that they were formed by a single item, and in the best case by 2, with very low weights; accordingly, we decided not to include them. The items excluded were as follows: (C=closed part; o =opened part): 3C, 8o, 10o, 14o, 21C, 23o, 24C and 25C. We established that the number of factors to be extracted was 5. Thus, we obtained 5 factors with eigenvalues greater than unity (see Table 3), which explained 52.18% of the variance. Table 3 shows the factors that belong to each of the 5 components obtained.

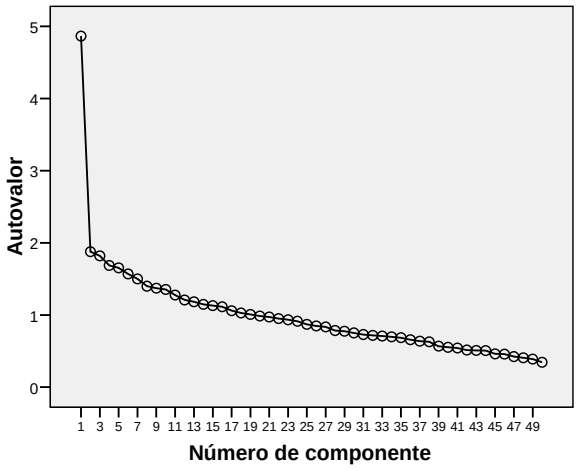

Figure 1: Figure showing the sedimentation obtained with the data.

Table 2: Variance explained of the refactoring of the first 14 factors obtained in the first F.A., forcing 5 factors.

	Sums of the squared saturations of the extraction		
Component	Total	% variance	% accumulated
1	2.515	15.720	15.720
2	1.787	11.168	26.888
3	1.599	9.993	36.882
4	1.348	8.425	45.307
5	1.100	6.873	52.180

Method of extraction: Principal components.

Table 3: Matrix of rotated components in the refactoring .

	Component				
	1	2	3	4	5
Factor 1	.774				
Factor 7	.772				
Factor 5	.597				
Factor 11		.751			
Factor 13		.612			
Factor 4		.575			
Factor 12			.768		
Factor 14			.684		
Factor 6				.714	
Factor 10				.584	
Factor 2				.581	
Factor 8		.404			-.730
Factor 3					.643
Factor 9			.420		.447

Method of extraction: Principal components.
Method of rotation: Promax normalization with the Kaiser Test.

Table 4 shows the items comprising the 5 factors obtained. As may be seen, they have little to do with Halpern's theoretical structure. Each of the factors is formed by items that belong to several of the Critical Thinking skills established by Halpern and not to only one of them, as would be expected. Also, there are some occasions on which the open part and the closed part of the same item –for example items 4o and 4C, item 5o and 5C, item 15o and 15C, items 12o and 12C, etc., form part of different factors, showing that they are measuring different aspects.

4.- Discussion

Regarding the reliability of the test, we have seen that it is quite low for the different skills included in it, with the exception of the reliability achieved with Hypothesis Testing, which is moderate. However, for the scale considered globally, this reliability reaches acceptable indices (0.77). Regarding the validity of the construct, in the component matrix, obtained in the Factor Analysis, we observed that not all the items contributed to a single factor and hence there a lack of unidimensionality. This can be considered as an important support for the multidimensional concept of Critical Thinking; that is, a manner of Critical Thinking formed by an entire set of skills. This is the opinion of most theoreticians in the field

(e.g. Ennis, 1987, 1996; Halpern, 1998, 2003a, 2003b; Facione et al., 1990; Swartz and Perkins, 1990), who state that critical thinking would be formed by a whole group of skills.

Table 4: Factors and items that belong to each of the five factors obtained in the refactoring procedure.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
F1	I.13 Arguments (O) I.13 Arguments (C) I. 12 Arguments (C) I. 4. Hypothesis (O)				
F 7	I.9 Verbal R.(O) I.9 Verbal R. (C)				
F5	I. 18 Likelihood(O) I. 18 Likelihood (C)				
F11		I. 10 Verbal R. (C) I. 15 Verbal R. (C) I. 5Hipothesis (O)			
F 13		I. 25 S. Problem (C) I. 21 S. Problem (C)			
F 4		I. 8 Verbal R. (C) I.3Hypothesis (O) I 4 Hypothesis (C) I 2Hypothesis (O)			
F12			I. 14 Arguments (C)		
F 14			I. 6 R. Verbal R. (C) I. 7 R. Verbal R. (O) I. 6. Verbal R.(O)		
F10				I. 24 S. Problems (O) I. 12 Arguments (O) I. 20 Likelihood (O)	
F 6				I. 22 S. Problems (C) I. 23 Arguments (C) I. 20 Likelihood (C) I 7 Verbal R.(C)	
F 2				I 11 Arguments (O) I 11 Arguments (C) I 1 Hypothesis (C)	
F 8					**I. 16 Likelihood (O)** **I. 17 Likelihood (O)** **I. 22 Likelihood (C)**
F 3					I 19 Likelihood (O) I 19 Likelihood (C) I 22 S. Problems (O) I 15 Arguments (C) I 1 Hypothesis (O)
F 9					I 17 Likelihood (C) I 15 Arguments (O) I 5 Hypothesis (C)

I= item; the corresponding number; C= Closed part; o= Open part. The items in bold weigh negatively on that factor.

Additionally, the factorial structure obtained does not seem to have much to do with the theoretical structure held by Halpern, who proposes a concept of Critical Thinking formed by 5 skills: Hypothesis Testing Skills, Argument Analysis Skills, Verbal Reasoning Skills,

Likelihood and Uncertainty Skills, and Decision Making and Problems Solving Skills. Nevertheless, as can be seen from table 4 the resulting factors are different. The first factor would be formed by items belonging to Argument Analysis, Verbal Reasoning, Hypothesis Testing and Likelihood and Uncertainty. The second factor includes items corresponding to the skills of Hypothesis Testing, Verbal Reasoning, Problem Solving and Likelihood and Uncertainty. The third factor would include elements of the Argument Analysis and Verbal Reasoning skills. The fourth factor, following the same trend, combines items from several dimensions, such as Problem Solving, Argument Analysis, Hypothesis Testing, and Verbal Reasoning. The fifth factor mainly comprises items of Likelihood, although some are linked to Problem Solving and Argumentation.

We analyzed the content of the items and the type of question posed in the different items in order to seek an answer to the factor grouping obtained. Let us now look at the items of the first factor; in brackets we note the theoretical dimension to which it belongs and then discuss its content with a view to finding something that links them together.

- In item 4o (Hypothesis Testing), the respondent is requested to formulate two questions to help decide between two ways of slimming.
- In items 9o and 9C (Verbal Reasoning), the respondent is requested to explain how he or she would act in a situation in which an ambiguously term is defined and operativized is presented. In the closed part the participant must choose the best option from among different alternatives.
- Item 12C (Argument Analysis) requests that the individual must evaluate whether a series of affirmations are a conclusion, a reason or a counterargument.
- Items 13o and 13 C (Argument Analysis). First, the participant must offer an opinion about the possibility that Universities might demand a new requisite- performing some kind of social service –in order to graduate. The participant must offer this opinion based on an argument. Second, he/she must assess whether a series of affirmations are a conclusion, a reason or a counterargument.
- Items 18o and 18C (Likelihood and Probability): the participant simply has to determine the equality of probability of winning a lottery number; choose yes or no, and explain why.

Upon analyzing the items forming this factor, it may be seen that they are mainly linked to the skill of the formulation and assessment of arguments. There are also items in which the participant is asked to determine whether certain expressions are reasons, conclusions, or counterarguments. It is striking that this first factor also contains items of the dimensions that Halpern calls Hypothesis Testing and Probability and Uncertainty. When we analyzed these items we have seen that there is a question in which the participant is requested to offer an explanation for his/her thinking. Explanation is a process that it tightly linked to argumentation, which would account for the weight of these items in the first factor.

The second factor comprises items that mainly belong to Hypothesis Testing and, although to a lesser extent, Verbal Reasoning, and Problem Solving. Let us now explore the content and type of responses or process demanded in an attempt to detect anything in common among the items.

- 5o (Hypothesis Testing): the participant must select the data that best support the affirmation and the data that best refute it.
- 2o (Hypothesis Testing): the participant is requested to formulate two questions to help us(?) make a decision.

- 3o (Hypothesis Testing). The participant is asked to state two changes to determine whether a given publicity campaign would work or not.
- 4C (Hypothesis Testing). The participant must assess the importance of a series of affirmations with a view to making a decision.
- 8C (Verbal Reasoning). The participant is asked to assess the quality of an analogy way of reasoning.
- 10C (Verbal Reasoning). The participant is asked to choose from among a series of affirmations whether they are reasonable criticisms of a type of behavior.
- 15C (Likelihood and Uncertainty). The participant is requested to evaluate whether a series of affirmations express an opinion or a reasoned argument.
- 21o and 25C (Problem Solving). The participant is asked to express a problem in two different ways and to offer solutions.

It may be seen that it is not easy to see common aspects among them.

The third factor is homogeneous, since the elements forming it mainly comprise items of Verbal Reasoning and Argument Analysis, such that there would be considerable consistency in its items. However, factors four and five are not clear either, and items of several skills belong to them. It is therefore not easy to see common elements in the items forming this factor. The heterogeneity of its content and questions hamper this.

Accordingly, it seems clear that critical thinking is not a single skill but involves several abilities. Another important fact to be taken into account is that the empirical structure obtained does not correspond to the theoretical structure offered by Halpern. Does this then mean that Halpern's categorization is incorrect? Not. Of course there are aspects that could be criticized, but it would also be possible to cast doubt on them in other categorizations. For example, Halpern discriminates between Argument Analysis Skills and Verbal Reasoning Skills whereas in fact it is not so easy to discriminate between them. According to Halpern, the skills of Argument Analysis include skills needed to identify the conclusion of an argument, assess the quality of the reasons adduced and to evaluate the global strong points of the argument. Within Verbal Reasoning, she includes the skills required for a person to defend him/ herself against persuasive techniques that are embedded in every-day language. In this sense, we understand that for this it would be necessary for the person to identify and assess the quality of the ideas, of the explanations, or of the arguments put forwards. Thus, one could well be speaking in terms of the same skill, related to the analysis and assessment of arguments, but in fact we might be referring to one that might depend on a more general process; namely, reasoning. In this sense, we understand that both are reasoning skills. One could say the same as regards Likelihood and Decision-Making. Halpern differentiates both skills but in our opinion, it is not possible to discriminate between Likelihood and Decision-Making, because to make decisions it is necessary to make analyze and assess the probability of different alternatives. Therefore, we consider that Likelihood and Uncertainty also form part of Decision Making. Such problems are not exclusive to this classification, but are the result of the simple fact of categorization, since categorizations are no more than artificial divisions of reality and as such they force it by establishing limits. And this is possible much more true in a process such as critical thinking, in which all its skills are profoundly overlapped. When one is trying to solve a problem or perform a task it is not easy to determine that a single skill of critical thinking is involved in its solution since several skills may be involved in it. In the Cornell Critical Thinking Test, Ennis et al. (1985) point out that there are certain items that can be assigned simultaneously to several of the dimensions of the test because they are measuring different things at the same time, which hinders the extraction

of pure factors in a factor analysis. This is what happened to us with the HCTAES. Thus, this seems to be a fairly common problem in tests about thinking. The factors are formed by items of several skills because when we make decisions at the same time we are reasoning about the advantages and disadvantages of the different alternatives and we are applying techniques and strategies involved in the calculation of probabilities and uncertainty. And when we are assessing an argument, we are also deciding and, also, we must choose which information is to be considered important and the basis for our reasoning. It therefore seems logical that it would be difficult to limit a single skill when we are solving a problem, making a decision or elaborating an argument. In light of all this several skills could be involved.

Accordingly, the difficulty involved in extracting pure factors or a coherent factor structure with a theoretical model derives, at least in part, from the difficulty in delimiting or extracting "in a pure way" which skill of critical thinking may be being used at a given moment, since thinking skills are tightly interrelated during daily cognitive though .In particular, there is a strong interrelationship between decision making and problem solving. Thus, making a decision is solving a problem, and indeed a very specific type of problem, a decision problem. The very process of problem-solving requires a skill for making decisions from among all the alternatives for solving it that have been generated. Additionally, reasoning is a crucial process for the proper fulfillment of decision-making and problem solving. Reasoning skills are fundamental for problem solving. People use their knowledge to decide what to do and to infer how to achieve their aims (Johnson-Laird and Shafir, 1993). Reasoning is involved in the definition of the problem, in the choice of the relevant information to be interpreted, and in its assessment. Also, during problem-solving we make continual inferences about the different alternatives for a solution and their outcomes. In particular, certain skills of inductive thinking, such as generalizations, hypothesis testing, analyses of causality and analogies, are essential for the resolution of most problems. Likewise, reasoning and decision -making are often interlinked in daily life. People (Johnson-Laird and Shafir, 1993) reason in order to make decisions and justify their choices to themselves and others; they reason to determine the consequences of their beliefs and hypothetical actions to be taken; they reason to develop their plans of action; they make decisions about which information should be given priority; they make decisions about which actions they should undertake and which information they should consider as the basis of their choices. There is enormous interdependence between reasoning and decision making, and these processes are continuously recursive. Shafir, Simonson and Tversky (1993) have reported evidence of the importance of reasoning in choice; in particular, they make an analysis of choice based on reasoning. This approach identifies several reasons and arguments that participate in and influence our decisions, and the authors explain choice as a balance between reasoning in favor of and against the different alternatives.

Thus, a possible explanation of the difficulties involved in extracting a factorial structure according to Halpern's theoretical model, and possibly any other theoretical model, stems from the difficulty in knowing which skill has been activated by a given task, since several of them may be involved. Bearing in mind that the actual construct of critical thinking is complex and that its nature hampers theoretical demarcation, it seems logical that the validity of a test aimed at measuring it would be weak. Nevertheless, it could be that the situations, items or problems proposed by the test are not really good and clarifying to state that a single skill would be required. In this sense it would be necessary to engage in further research to determine the type of problems and the thought processes required to solve them. Recently, Saíz, Rodríguez and Fernández (in prep.) have developed a test with open questions; some of the bases are similar to those of Halpern's test, but the main difference lies

in the fact that the problems can only be solved in one way; that is, there is only one correct answer, whereas in the HCTAES there are several correct answers. However, in this test (Saiz et al. in prep.) the factor analyses performed are not coherent with the theoretical structure underlying the test either, even though they were grouped following other criteria -for example, general vs. specific problems- which at least allowed a certain coherence to be given to the factors extracted. Thus, the extraction of pure factors seems to be an arduous task in such a complex construct and one that is not very well developed in theoretical terms. For the time being, we believe that Halpern's test is a good one for assessing critical thinking because we have noted that its reliability indices are acceptable, indicating that it could be a good instrument for evaluating critical thinking. Nevertheless, it would be necessary to continue along this line of enquiry to corroborate our results and to analyze other aspects of its validity, such as convergent and divergent validity.

6. - References

APA (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Executive Summary "The Delphi Report". http://www.insightassessment.com/dex.html.

Delgado, A.M., Borge, R., García, J., Oliver, R. & Salomón, L. (2005). Competencia y diseño de la evaluación continúa y final en el Espacio Europeo de Educación Superior. http://www.usal.es/~ofeees/ARTICULOS/competencias_evaluacion_eees_mec_pdf

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. En J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills* (pp. 9-26). New York: Freeman and Company.

Ennis, R. H. (1996). *Critical thinking*. Upper Saddle River, NJ: Prentice-Hall.

Ennis, R. H. (2003). Critical thinking Assessment. In D. Fasko (Ed.) *Critical* thinking and Reasoning. Cresskill, NJ: Hampton Press.

Ennis, R. H., & Millman, J. (1985). *Cornell Critical Thinking Test, Level Z.* (3rd Ed.). Pacific Grove, CA: Midwest Publications.

Ennis, R. H., Millman, J. y Tomko, T. N. (1985). *Cornell Critical Thinking Tests LevelX & Level Z.* –Manual (3rd Ed.). Pacific Grove, CA: Midwest Publications.

Ennis, R. H., & Weir, E. (1985). *The Ennis-Weir Critical Thinking Essay Test.* Pacific Grove: Midwest Publications.

Facione, P. et al. (1990). *The California Critical Thinking Skills Test (CCTST): Forms A and B; and the CCTST test manual.* Millbrae. CA: The California Academic Press.

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*(4), 449-455.

Halpern, D. F. (2003a). *Thought and knowledge. An introduction to critical thinking* (4ª Ed.). Hillsdale, NJ: Erlbaum.

Halpern, D. (2003b). Halpern Critical Thinking Assessment Using Everyday Situations: Background and scoring standards. Unpublished manuscript. Claremont, CA: Claremont McKenna College.

Halpern, D. (2006). Halpern Critical Thinking Assessment Using Everyday Situations: Background and scoring standards (2º Report). Unpublished manuscript. Claremont, CA: Claremont McKenna College.

Jacobs, S. S. (1995). Technical characteristics and some correlates of the California Critical Thinking Skills Test Forms A and B. *Higher Education Research, 36*, 89-108.

Johnson-Laird, P. N., & Shafir, E. (1993). The interaction between reasoning and decision making: An introduction. *Cognition, 49*, 1-9.

Loo, R. & Torpe, K. (1999). A psychometric investigation of scores on the Watson-Glaser Critical Thinking Appraisal New Form S. *Educational and Psychological Measurement, 59* (6), 995-1003

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition, 49*, 11-36.

Swartz, R. J., & Perkins, D. N. (1990). *Teaching thinking: Issues and approaches*. Pacific Grove, CA: Critical Thinking Press & Software.

Watson, G., & Glaser, E. M. (1984). *The Watson-Glaser Critical Thinking Appraisal.* San Antonio, TX: The Psychological Corporation.