Journal of Education and Human Development
March 2016, Vol. 5, No. 1, pp. 240-248
ISSN: 2334-296X (Print), 2334-2978 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/jehd.v5n1a24
URL: http://dx.doi.org/10.15640/jehd.v5n1a24

The Effects of Teaching Critical Thinking Persist Over Time

Silvia. F. Rivas & Carlos Saiz[1]

Summary

The aim of the work is to determine whether the skills of critical thinking persist over time; i.e., whether they endure after training. To accomplish this, we carried out a follow-up study of a sample instructed in critical thinking four years previously. We applied an instruction program, ARDESOS, to this sample and assessed the improvement in the basic skills of critical thinking. After checking the improvement in competencies we re-appraised it four years later. The results show that once the students' university studies had been completed these competencies persisted and even improved with respect to those acquired after instruction in all the dimensions of critical thinking except one. Here we detail the whole process of evaluation and the results obtained.

Keywords: Critical thinking, problem solving, decision making, evaluation

1. Introduction

As a consequence of the new realities and conditions in which current society is now immersed, where knowledge flows at an ever speedier rate and technological advances are unceasing, the competencies required in our professional activities are also becoming increasingly more important, and education is an indispensable tool for adapting to such changesthis situation means that the ability to think critically is more important than ever in order to achieve a minimum of personal wellbeing and adequate competency in any context (Saiz, 2015). Within the sphere of higher education, in many countries these transferrable competencies in critical thinking are being incorporated as objectives in University curricula. However, less interest has been focused on the assessment of these skills to see whether they really are good or not and whether they persist along the years of study. The work presented here addresses this issue.

To improve critical thinking skills, it is important to evaluate not only the effects of instruction program, but also its efficiency, persistence and transference. Efficiency is determined by the ability of a given teaching initiative to improve thinking skills. To evaluate efficiency, we must assess whether students' performance improves after receiving instruction as compared with the time before the intervention; i.e., we must evaluate efficiency as a function of the size of the effect of the intervention. Research in this field has mainly focused on addressing whether instruction programs are efficient and hence whether students improve after receiving such instruction. However, many instruction initiatives report results on efficiency alone. To cite some of the more classic works, Perkins and Grotzer (1997) made an exhaustive review of instruction initiatives of general programs such as "Philosophy for Children" by Lipman; "Productive Thinking" by Covington et al., or "Instrumental Enrichment" by Feuerstein, with positive results, showing the efficiency of these programs in improving the skills of critical thinking. Along this line, there are also extensive reviews of programs designed to teach how to think that can be consulted by interested readers (Maclure & Davies, 1991/1994; Nickerson, Perkins & Smith, 1985/1987; Saiz, 2002; Segal, Chipman & Glaser, 1985). Halpern (2014) offered a more updated review of studies aimed specifically at addressing the different mechanisms of critical thinking, all of them also with favourable evaluations after instruction.

[1] University of Salamanca (Spain), Facultad de Psicología, Avda. de la Merced, 109-131, 37005 Salamanca, Spain.

In all the above studies, the authors evaluate the efficiency of the programs as a function of the *effect size* of the intervention and, in many of them, also the *transference* of the skills to new contexts and areas of knowledge. However, this should not be considered sufficient and it would be desirable that the changes or improvements achieved should persist over time; i.e., the effects of teaching should endure beyond the end of the instruction. Here we are referring to the degree of *persistence* (or *permanence*) of the effect of learning. Measuring persistence is an undeniable commitment, as it requires longitudinal studies. Therefore, studies addressing this aspect are scarce. Perkins and Grotzer (1997), concluded that in most of these general programs alluded to previously the persistence of the effects are not measured. Only Adey and Shayer's (1993) "Thinking in Science" program (CASE; Cognitive Acceleration through Science Education) and Lipman's "Philosophy for Children" program (Lipman, Sharp & Oscanyan, 1980) offer positive data concerning the persistence of the skills taught in both programs. Adey and Shayer (1993) made three measurements of their instruction program: one at the end of the program and the other two 1 and 2 years later. The results revealed improvements in critical thinking skills as a result of the instruction and their persistence and improvement up to two years after the program had been completed. Along the same lines, data pointing to persistence were obtained two and a half years later in the case of the "Philosophy for Children" program. In a recent study (Arum & Roksa, 2011), the authors followed up more than 2,300 students from 24 universities over four years to assess whether students were learning the specific skills of critical thinking during their university studies. The results of this study were not very promising since they showed that the skills did not improve even 3 years after the students' university studies had been completed. They observed that 40% of the students did not improve their critical thinking skills appreciably during the first two years of university studies. After 4 years, 36% did not show significant advances in these skills either. The authors considered that such findings could be relevant for focusing attention on the quality of the education received by university students. However, and despite being an important reference (since there are no others with these characteristics in the literature), their study did not aim to measure the persistence of critical thinking skills after an intervention, it only monitored the students' critical thinking skills over several years. Nevertheless, since it provides an assessment of the competencies of critical thinking in university studies, we consider it to be an important contribution. However, the following should be borne in mind: according this work, after years of study in classrooms, the critical thinking skills of the students declined, which is the opposite of what would be expected. Despite the relevance of the data in their study, it is not relevant for the analysis of persistence, since it did not apply an instruction program.

Bearing in mind this gap, for some years we have been refining our own program ARDESOS (from the Spanish, equivalent to **A**rgumentation, **De**cision, **So**lving of problems in daily **S**ituations) for the improvement of critical thinking skills and evaluating its efficiency via the PENCRISAL standardized test (from the Spanish, equivalent to Critical Thinking Salamanca –**Pen**samiento **Crí**tico **Sal**amanca). The instruction program, improved twice, has provided increasingly satisfying results, as may be seen in Saiz and Rivas (2011) and Saiz, Rivas and Olivares (2015). However, as highlighted above, evaluating the efficiency of an instruction program is not sufficient: transference and persistence are also important elements.

Since our instruction program has been tested on psychology undergraduates, after three years we now have the opportunity of evaluating whether in the students' fourth year these critical thinking skills, which improved in their first year at university, persist after three years. Measuring whether these changes persist after three academic years would be extremely important, and it is indeed the aim of the present study. It is necessary to ensure that the new thinking skills acquired after an intervention will persist and do not fade over time, when the learning content is not present. This should contribute data regarding two issues: on one hand, that the ARDESOS program for improving critical thinking skills is efficient as regards the persistence of such skills and, on the other hand, that if such skills increase this could indicate that the program really is favouring the development of horizontal competences in the degree course, one of the challenges proposed in the context of the European Higher Education Area (EHEA).

Thus, the main aim of the present work is to determine to what extent the skills of critical thinking increase and whether such improvements persist for some time after the intervention. To accomplish this, we measured performance in critical thinking skills when the students had completed their degree in Psychology and compared the results with those obtained after application of the ARDESOS program in the first year. Accordingly, there were two sub-aims: 1) to demonstrate if skills are indeed acquired after the intervention and 2) to check whether such skills persist over time.

Method

Participants

The sample used for this study was made up of 81 students in the final year of the degree in psychology offered by the University of Salamanca (Spain). Of these, 86.4% (70) were women and the rest (13.6%) men, such that –statistically-the majority were women with $p<.001$ ($Chi^2= 42.975$; 1 df; $p=.000$). The mean age of the sample was 21.21 (S.D.=.904) (IC at 95%: 21.01-21.41), in a range of 20-26 years, with a distribution that was not normal, with $p<.01$ in the Kolmogorov-Smirnov test of goodness of fit ($p=.000$), owing to a marked positive asymmetry (As: 2.998) and a clearly leptokurtic curve (K=11.673).

Program for the development of the ARDESOS critical thinking test.

The aim of our intervention (ARDESOS program) is to optimize the intellectual skills involved in critical thinking; that is, *Reasoning, Problem-solving* and *Decision-making* through the use of daily situations for their development.

Our proposal is designed for application in the classroom in 50-60 hours of physical presence there, distributed in some 15 sessions of weekly work lasting 3-4 hours each and covering a total time of 120-140 hours (60 hours of physical presence plus 60-80 hours of students working on their own).

The ARDESOS program is based on the following principles:

- It is a *direct method for the teaching* of thinking skills, since this type of teaching allows the transfer of knowledge to be achieved.
- The program essentially places emphasis on the *learning of processes*, in view of the mainly procedural nature of these capacities.
- It is based on teaching-learning strategies of *Problem-Based Learning* (PBL), where activity revolves around the discussion of the different situational problems designed in the program and the learning of the skills of critical thinking arises from the experience of having worked on such situations.

For further information about our program and the results of its efficiency, the following works may be of help (Olivares, Saiz & Rivas, 2013; Saiz & Rivas, 2008a, 2011, 2012; Saiz, Rivas & Olivares, 2015).

Instruments

The PENCRISAL critical thinking test.

In order to measure the size of the effect after an intervention program (ARDESOS) and the persistence of the students' thinking skills along their studies in psychology, we applied the PENCRISAL critical thinking test developed by Saiz and Rivas (Rivas & Saiz, 2012; Saiz & Rivas, 2008b). This test has been used in different studies carried out in our Department (Butler, Dwyer, Hogan, Franco, Rivas, Saiz, & Almeida, 2012; Olivares, Saiz & Rivas, 2013; Saiz & Rivas, 2011, 2012; Saiz, Rivas & Olivares, 2015).

PENCRISAL is a battery comprising 35 situational problems with an open response format, with different thematic problems of knowledge and single responses. The problem situations are distributed in *5 factors* that make up the basic skills of thinking, and within each of these we find the forms of reflection and solution most relevant to our daily lives, as described below (Cronbach Alpha= .632; test-retest: r=.786, Rivas & Saiz, 2012).

Deductive reasoning: Evaluation of propositional reasoning and categorical reasoning.

Inductive reasoning: Evaluation of analogical reasoning, hypothetical reasoning and inductive generalizations.

Practical reasoning: Evaluation of argumentation skills and the identification of fallacies.

Decision-making: Evaluation of the use of general decision procedures, which involves the elaboration of precise judgements of probability and the use of suitable heuristics to make sound choices.

Problem-solving: Evaluation of the setting up of specific strategies for solving the situational problems posed.

PENCRISAL was applied through the Internet using the following evaluation platform: SelectSurvey.NET: http://survey.pensamiento-critico.com/Login.aspx

The test provides a total score of the critical thinking skills and another five scores that refer to the 5 factors: Deductive Reasoning (DR), Inductive Reasoning (IR), Practical Reasoning (PR), Decision-making (DM) and Problem-solving (PS). The range of values lay between 0 and 70 points maximum for the overall score of the test and 0-14 for each of the five scales.

Procedure

The study was carried out from September to January 2014-2015 with students from the fourth year of the Degree in Psychology offered at the University of Salamanca. This group of students had already undergone an initial measurement (pre-treatment) of their critical thinking skills when they began their degree studies (2011-2012). Also, all the students received an instruction program for improving these skills through the ARDESOS program and another post-treatment measurement was made to check the efficiency of the program.

At the beginning of the 2014-2015 academic year all the students registered in the fourth year (about 220) were invited to collaborate voluntarily in this project. Emphasis was placed on the interest of this type of study aimed at investigating the changes in critical thinking skills that occurred during their university years and that would allow us to know to which extent the instruction program (ARDESOS) they followed in their first year or the actual training received for their degree had improved these competencies or not. This would allow us to make an assessment and propose improvements, if necessary. Accordingly, we measured the competencies in thinking at the end of the degree studies, also via the PENCRISAL test.

As explained above, the time elapsed between the first measurement (pre-treatment) and the second one (post-treatment) was 4 months. The time between the post-treatment measurement and the measurement of permanence at the end of the degree was 3 years.

Design

In order to analyze the effect of the intervention and the persistence of its effects, a quasi-experimental longitudinal study with pre-post measurements was designed.

Statistical analyses

For the statistical analyses the IBM SPSS Statistics 21 package was used. The following tools and statistical techniques were used: frequency tables and percentages for qualitative variables, with the Chi-squared homogeneity test; exploratory and descriptive analyses for quantitative variables, with goodness fit test to the normal Gaussian model and box diagrams for the detection of outliers; the usual descriptive statistics (mean, standard deviation, median, etc.) in numerical variables; Student's t tests for the significance of differences of means. To estimate effect size, we used Cohen's d statistic (1988) for related groups.

Results

Regarding the descriptive statistics of all the variables studied, we observed that only pre-treatment problem-solving did not fit the normal curve ($p > .050$) and that post-treatment decision-making showed a slight deviation from the norm, although the p values was still $<.010$. The other variables fitted the normal curve adequately, with $p < .050$. Below we offer a summary table of the descriptive statistics of the variables explored in the study.

Table 1: Descriptive statistics of the PENCRISAL variable

Variables	n	M	SD	95% CI	Range (min-max)	K-S p-sig
Pencrisal TOTAL-Pre-treatment	81	28.49	6.201	[27.12, 29.86]	12-43	.531
Deductive-Pre- treatment	81	4.14	1.941	[3.71, 4.57]	0-9	.140
Inductive-Pre- treatment	81	5.07	1.523	[4.74, 5.41]	2-9	.075
Practical-Pre- treatment	81	6.10	2.364	[5.58, 6.62]	0-11	.130
Decision-making-Pre- treatment	81	6.54	1.761	[6.15, 6.93]	2-11	.197
Problem-solving- Pre - treatment	81	6.16	2.094	[5.70, 6.62]	1-10	.001
Pencrisal TOTAL-Post- treatment	81	31.44	6.068	[30.10, 32.79]	19-46	.603
Deductive- Post - treatment	81	5.27	2.162	[4.79, 5.758]	0-11	.292
Inductive- Post - treatment	81	5.47	1.718	[5.09, 5.85]	2-9	.061
Practical- Post - treatment	81	8.25	1.978	[7.81, 8.68]	4-13	.131
Decision-making- Post - treatment	81	7.09	1.976	[6.65, 7.52]	3-13	.012
Problem solving- Post - treatment	81	5.37	2.266	[4.87, 5.87]	0-11	.242
Pencrisal TOTAL-Persistence	81	35.48	5.211	[34.33, 36.63]	23-48	.803
Deductive- Persistence	81	4.53	1.911	[4.11, 4.95]	0-8	.230
Inductive- Persistence	81	6.01	1.270	[5.73, 6.29]	3-9	.025
Practical- Persistence	81	9.56	2.716	[8.96, 10.16]	3-14	.104
Decision-making- Persistence	81	8.36	1.958	[7.93, 8.79]	4-12	.281
Problem-solving- Persistence	81	7.02	2.236	[6.53, 7.52]	1-11	.116

With a view to assessing the persistence of critical thinking skills over time, we first explored the results concerning the efficiency of the program, after which we checked the effect of the passage of time on performance in critical thinking acquired after the intervention. To accomplish this, we measured the significance of the differences between means using Student's t-tests. To add quantitative information from the dimension of the impact of the intervention, we also calculated the effect size with Cohen's d (1988) to determine the standardized difference of means.

As may be seen in table 2, the results show that application of the ARDESOS program reveals significant differences between pre- and post- performance right across the scale ($t_{(80)}$=-4.195, p=.000) and in the factors Deduction ($t_{(80)}$=-4.391, p=.000), Induction ($t_{(80)}$=-2.100, p=.019), Practical Reasoning ($t_{(80)}$=-7.596, p=.000) and Decision making ($t_{(80)}$=-2.061, p=.021). In all these it may be seen that the means increase as a result of the intervention (M_{TOTpre}=28.49; $M_{TOTpost}$=31.44); (M_{DRpre}=4.14; M_{DRpost}=5.27); (M_{IRpre}=5.07; M_{IRpost}=5.47); (M_{PRpre}=6.10; M_{PRpost}=8.25) and (M_{DMpre}=6.54; M_{DMpost}=7.09). However, this was not the case for Problem-Solving: although the difference between the two means was statistically significant ($t_{(80)}$=2.694, p=.004), the participants showed poorer performance after the instruction (M_{PSpre}=6.16; M_{PSpost}=5.37). The results show that the ARDESOS program is effective for most of the skills of critical thinking since performance improved significantly in all the factors with the exception of Problem-Solving.

Table 2: Comparison of the pre-post-treatment means of PENCRISAL

Variables		n	M	SD	Difference between means 95% IC	Sudent's t test T value	df	p-sig (Unilateral)	Effect size (Cohen's d)
Total	Pre	81	28.49	6.201	-2.951 [-4.350, -1.551]	-4.195	80	.000**	0.49
	Post	81	31.44	6.068					
Deduction	Pre	81	4.14	1.941	-1.136 [-1.651, -.621]	-4.391	80	.000**	0.53
	Post	81	5.27	2.162					
Induction	Pre	81	5.07	1.523	-0.395 [-.769, -.021]	-2.100	80	.019*	0.23
	Post	81	5.47	1.718					
Practical	Pre	81	6.10	2.364	-2.148 [-2.711, -1.585]	-7.596	80	.000**	1.09
	Post	81	8.25	1.978					
Decision-making	Pre	81	6.54	1.761	-.543 [-1.068, -.019]	-2.061	80	.021*	0.28
	Post	81	7.09	1.976					
Problem- solving	Pre	81	6.16	2.094	.790 [.206, 1.167]	2.694	80	.004**	0.35
	Post	81	5.37	2.266					

* Significant at 5% **Significant at 1%

These findings are consistent with those observed in our earlier study in which we reported the results of the efficiency of version 2 (v2) of the ARDESOS program. (Saiz, et al., 2015). The improved ARDESOS v2 program shows effect sizes very similar to those observed in the present study. In ARDESOS v2, the effect size of Practical Reasoning (d=0.83) and Deductive Reasoning (d=0.63) takes very high values. Likewise, the total of the scale (d=0.48) and the Problem-Solving factor (d=0.44) have a moderate effect size. Those with the smallest effect size are Inductive Reasoning (d=0.23) and Decision-making (d=0.16). This same pattern appears in the present study, where the factors with the highest values in effect size are Practical Reasoning (d=1.09) and Deduction (d=0.53). The total of the scale (d=0.49) has moderate values and, likewise, those with the lowest values are Induction (d=0.23) and Decision-Making (d=0.28).

Having seen that the ARDESOS program afforded an improvement in performance in critical thinking skills, we now describe the results concerning persistence to check whether the skills were maintained over time.
The results (see table 3) point to significant differences between the two measurements in the total of the PENCRISAL test ($t_{(80)}$=-6.504, p=.000) and in Induction ($t_{(80)}$=-2.533, p=.013), Practical Reasoning ($t_{(80)}$=-4.431, p=.000), Decision-making ($t_{(80)}$=-4.783, p=.000) and Problem-solving ($t_{(80)}$=-4.983, p=.000). These differences are reflected in the means since the scores on performance increase in the second measurement (persistence) ($M_{TOTpost}$=31.44; $M_{TOTpersist}$=35.48); (M_{IRpost}=5.47; $M_{IRpersist}$=6.01) (M_{PRpost}=8.25; $M_{PRpersist}$=9.56); (M_{DMpost}=7.09; $M_{DMpersist}$=8.36); (M_{PSpost}=5.37; $M_{PSpersist}$=7.02).This indicates that not only that the critical thinking skills were maintained 3 years after the ARDESOS intervention but also that they improved, since performance increased significantly with respect to the that acquired in the intervention. Nevertheless, in deduction the significant difference between the two measurements ($t_{(80)}$=2.572, p=.007) runs in the other direction; that is, performance worsened over time (M_{DRpost}=5.27; $M_{DRpersist}$=4.53).

Table 3: Comparison of post-treatment means-persistence of PENCRISAL

Variables		n	M	SD	Difference between means (IC 95%)	Student's t test T value	df	p-sig (Bilateral)	Effect size (Cohen's d)
Total	Post	81	31.44	6.068	-4.037 [-5.272, -2.802]	-6.504	80	.000**	0.78
	Persist	81	35.48	5.211					
Deduction	Post	81	5.27	2.162	.741 [.205, 1.276]	2.752	80	.007**	0.40
	Persist	81	4.53	1.911					
Induction	Post	81	5.47	1.718	-.543 [-.970, -.116]	-2.533	80	.013*	0.43
	Persist	81	6,01	1.270					
Practical Reasoning	Post	81	8.25	1.978	-1.309 [-1.896, -.721]	-4.431	80	.000**	0.48
	Persist	81	9.56	2.716					
Decision-making	Post	81	7.09	1.976	-1.272 [-1.801, -.743]	-4.783	80	.000**	0.65
	Persist	81	8.36	1.958					
Problem-solving	Post	81	5.37	2.266	-1.654 [-2.315, -.994]	-4.983	80	.000**	0.74
	Persist	81	7.02	2.236					

* Significant at 5% **Significant at 1%

Finally, we feel it to be of interest to compare the effect size of the efficiency of the program described above, with those obtained for the persistence measurements. Our data indicate that in the measurement of persistence a noteworthy increase can be seen in the effect size of the overall performance of the scale (d=0.78) with respect to that of post-treatment (d=0.49). The same is the case for induction (d=0.43 for persistence and d=0.23 for post treatment), Decision-making (d=0.65 for persistence and d=0.28 for post-treatment), and Problem-solving, which did not undergo improvements even with the intervention, (d=0.74 for persistence and d=0.35 for post-treatment).

In light of these results, it may be concluded that the changes introduced improve overall critical thinking skills, albeit less so for problem-solving. Moreover, these skills improve over time, with the exception of deductive reasoning. In comparative terms, the skills with the greatest effect size, after the program, are practical reasoning and deduction, and to a lesser extent decision-making and induction. Finally, regarding persistence, the skills of problem-solving and decision-making are best maintained over time, followed by practical reasoning and induction.

Discussion

Overall, the results of the present study show that the ARDESOS program for developing the skills of critical thinking meets the main criteria of efficiency proposed by Perkins and Grotzer (1997). On one hand, our findings show that the ARDESOS program works, since it improved the performance of the students who followed it. However, as already pointed out, success should not only be expressed in terms of the criterion of effect size; instead, this effect should also be apparent in tasks different from those of the learning context and these newly acquired skills should persist over time. This criterion -the degree of persistence of the effect of learning- is also seen in the present study. Moreover, the students did not only retain the skills acquired as a consequence of the ARDESOS instruction; these skills increased considerably in the dimensions of Inductive Reasoning, Practical Reasoning, Decision-making and Problem-solving over the 3-year degree period.

Specifically regarding effect size, our results indicate that there are significant differences with respect to critical thinking skills between the pre- and post-treatment measurements in the total of the scale and in the Deductive Reasoning, Inductive Reasoning, Practical Reasoning and Decision-making factors. It should be noted that this was not the case of Problem-solving skills, which worsened after the intervention. This is perhaps due to the nature of the items, in particular those referring to general strategies. In the instruction, these skills were worked globally, not specifically. The lack of specific practice may be the reason underlying this lack of efficiency in this dimension. This problem is addressed in version 3 of our program, currently under development.

The results obtained here are consistent with those of our previous study in which we evaluated the efficiency of the program used here (Saiz, et al., 2015), with very similar values as regards effect size in both studies. Likewise, the data reported here allow us to conclude that once the program has been completed the competencies acquired by the students do not only persist, as was to be expected, but are also improved. All the students who followed the ARDESOS program improved their skills on the total of the scale and in Inductive and Practical Reasoning, Decision-making and Problem-solving three years later. The only skill that did not persist, but worsened, was Deductive Reasoning.We believe that the lack of improvement in deductive reasoning is not so much due to any intrinsic difficulty but rather to the difficulty in applying it to everyday situations. Inductive mechanisms are easier to apply because humans (and mammals in general) are above all part of an inductive machinery owing to the simple principle of species conservation. For example, identifying regularities is a difficult inductive mechanism but it is easy for humans (and mammals) because they do it from a very early age because of its vital importance in survival. Thus, the passage of time penalizes deduction more than it does induction.

Persistence is another relevant indicator of the efficiency of an instruction program. We were saying that the measurement of persistence has been somewhat overlooked and only the programs of Adey and Shayer, and the "Philosophy for Children" program measured persistence after several years, with positive results. By contrast, a study by Arum and Roksa has reported rather unpromising results, although in this case there was no instruction program, and the assessment only looked at the changes produced by the study plan itself. By contrast, the study of Arum and Roksa cited above reports data in which critical thinking competences decline over the years. As we have stated, the problem with their study's data is that it lacked instruction, and therefore, they are not relevant at this point, although they are worrying.

In our case, regarding the effect size of the change in persistence, it should be noted that here the skills in first place would be Problem-solving, Decision-making, Practical Reasoning, and Inductive Reasoning.The fact that the critical thinking skills that improve most are the first two can be better understood taking into account the following remarks. One of the main aspects of critical thinking refers to mechanisms of inference, although these are not the only ones; there are others that are mainly action strategies, determined by other processes in essentially inductive processes. Actually, these mechanisms are responsible for the real change of events and these are the fruit of daily, ongoing experience. It is thus possible to understand why the mechanisms of problem solving and decision making improve and persist longer than the rest. In short, they are the processes most often used in daily life, such that they benefit more from instruction than others less used, such as deductive processes. Moreover, experience and daily academic practice and the knowledge gained in the degree course itself may also contribute to the improvement of these skills over time.

In both their academic and daily lives, along their university years' students evaluate or produce arguments to defend points of view; they use analogical and causal arguments, and they seek explanations. All this has a final objective: that of making decisions aimed at solving problems as efficiently as possible.

Another important aspect that could support this persistence and improvement in the skills of critical thinking is transference. Instruction programs must be designed to achieve a generalization of critical thinking skills to different situations and areas of knowledge. Research into the teaching and learning of critical thinking has shown that for instruction in critical thinking to be effective and for transference to occur, two key aspects must be taken into account, namely, that the skills of critical thinking should be taught by direct instruction and that these skills should be taught for a broad range of contexts that simulate reality (Halpern, 1998; Marin & Halpern, 2011). Our ARDESOS program meets these requirements since it encourages students to use the skills acquired in instruction in general contexts of the type they are likely to meet in everyday life. Moreover, the repeated use of these skills probably contributes to their improvement.

It is also important to take into account that one is dealing with a test whose actual administration is difficult. Such a test is necessary in this type of instrument since only in this way would it be possible to demonstrate the effect of the intervention, without having to design another parallel instrument for the purpose. Accordingly, is it not only important that we found differences in performance each time the measurements were made but also that, despite the considerable difficulty inherent to its items, the improvement in performance reached such high scores. After the program, performance increased to the 70th percentile and after three years it reached 85%. These important differences indicate that one is not dealing with just any type of improvement but, instead, with a noteworthy increase in competencies.

Despite the foregoing, we are aware that our study has a series of limitations that could hamper the generalization of the results. Examples are the lack of a control group whose members had not received the ARDESOS program in order to be able to check whether the skills had persisted after 3 years or not and, also, the influence of variables such as the age and the level of maturity of the students. An explanation of why the students increased their skills so much along their degree course could lie in the variables just mentioned. Accordingly, it would be interesting to carry out further studies to check this. The sample we were working with comprised students from the same academic year, such that age and maturity were the same in the measurement of persistence. It would be interesting, therefore, to start out with an initial sample with a broad age range to be able to see, in the final measurement, the performance in critical thinking at those ages. The same can be said of other degrees, such as in the sciences and humanities, to assess whether all these students evolve to the same extent, since it is possible that humanities students might not develop as much as science students owing to the specific natures of their respective curricula. These degree courses demand the use and development of different skills, and it is foreseeable that there would be a different pattern in the improvement of the different competencies as a result of the different curricular contents of each of the degree courses.

In conclusion, we feel our results are highly satisfactory as regards the performance of our students. As discussed, we believe it is important to assess persistence after applying an instruction program. We believe that this is where we should be concentrating our efforts – in carrying out studies and providing evidence regarding the measurement of *persistence,* since this is a contribution to the whole of the sphere of teaching and, in our particular case, of the teaching of competencies in critical thinking. In the future, we hope to delve further into this line of enquiry, which should allow us to analyze more precisely whether the development of horizontal competencies from university degrees is really being fostered.

References

Adey, P., & Shayer, M. (1993). An exploration of long-term far-transfer effects following an extended intervention program me in the high school science curriculum. *Cognition and Instruction, 11*(1), 1 - 29. doi: 10.1207/s1532690xci1101_1

Arum, R., & Roksa, J. (2011). *Academically Adrift: Limited Learning on College Campuses.* Chicago: The University of Chicago Press.

Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., & Almeida, L. F. (2012). Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications**.** *Thinking Skills and Creativity, 7*, 112-121. doi: S1871187112000247

Cohen, J. (1988). *Statistical power analysis for behavioral sciences.* (2nd ed.). Hillsdale, NJ: Erlbaum.

Halpern, D.F. (1998). Teaching critical thinking for transfer across domains - Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist, 53* (4), 449-455. doi: 10.1037/0003-066X.53.4.449

Halpern, D. F. (2014). *Thought and Knowledge. An Introduction to Critical Thinking* (5th Edition). New York: Psychology Press.

Lipman, M., Sharp, A., & Oscanyan, F. (1980). Philosophy in the classroom. *Philadelphia: Temple University Press.*

Maclure, S., & Davies, P. (Eds.). (1994). *Aprender a pensar, pensar en aprender.* Barcelona: Gedisa. (Original 1991).

Marin, L. M., & Halpern, D.F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity, 6,* 1-13. doi: S1871187110000313

Nickerson, R.S., Perkins D.N., & Smith E.E. (1987). *Enseñar a pensar. Aspectos de la aptitud intelectual.* Madrid: Paidós/M.E.C. (trad. cast.: L. Romano y C. Ginard. Original de 1985).

Olivares, S., Saiz, C., & Rivas, S.F. (2013). Encouragement for thinking critically. *Electronic Journal of Research in Educational Psychology, 11* (2), 367-394.doi:10.14204/ejrep.30.12168

Perkins, D. N., & Grotzer, T.A. (1997). Teaching intelligence. *American Psychologist, 52* (10), 1125-1133.

Rivas, S.F., & Saiz, C. (2012). Validación y propiedades psicométricas de la prueba de pensamiento crítico PENCRISAL. *Revista Electrónica de Metodología Aplicada*, *17* (1), 18-34.
http://www.pensamiento-critico.com/archivos/validacionpencrieng.pdf

Saiz, C. (2002). Enseñar o aprender a pensar. *Escritos de Psicología, 6,* 53-72.

Saiz, C. (2015). Efficay, the heart of critical thinking. En Domínguez, C. (ed.). *Pensamento crítico na educação: Desafios atuais. (Critical thinking in education: Actual challenges)*. Pp. 159-168. Vila Real: UTAD.
http://www.pensamiento-critico.com/archivos/ebookutad.pdf

Saiz, C., & Rivas, S.F. (2008a). Intervenir para transferir en pensamiento crítico. *Praxis. 10* (13), 129-149.

Saiz, C., & Rivas, S.F. (2008b). Evaluación del pensamiento crítico: una propuesta para diferenciar formas de pensar. *Ergo, Nueva Época, 22-23*, 25-66.
http://www.pensamiento-critico.com/archivos/evaluationCTergoENGLSH.pdf

Saiz, C., & Rivas, S. F. (2011). Evaluation of the ARDESOS program: An initiative to improve critical thinking skills. *Journal of the Scholarship of Teaching and Learning, 11* (2), 34-51.

Saiz, C., & Rivas, S.F. (2012). Pensamiento crítico y aprendizaje basado en problemas. *Revista de Docencia Universitaria, 10* (3), 325-346.

Saiz, C., Rivas, S.F., & Olivares, S. (2015). Collaborative learning supported by rubrics improves critical thinking. *Journal of the Scholarship of Teaching and Learning. 15* (1), 10-19.doi: 10.14434/josotl.v15i1.12905

Segal, J. W., Chipman, S. F., & Glaser, R. (1985) *Thinking and learning skills. Volume. 1: Relating instruction to research*. New York: Erlbaum.