Electronic Journal of Research in Educational Psychology
ejrep
EDUCATION & PSYCHOLOGY I+D+i
Ilustre Colegio Oficial de Psicología de Andalucía Oriental

Psychometric properties of the Scientific Inquiry Scale

Carlos Ossa-Cornejo[1], Alejandro Díaz-Mujica [2], Jaime Aedo-Saravia [3], Jose M. Merino-Escobar[4], Claudio Bustos-Navarrete[2]

[1] Dept. of Educacional Science, University of Bio-Bio
[2] Dept. of Social Science, University of Concepcion
[3] Dept. of Social Science and Communication, University Santo Tomas
[4] Dept. of Education, University of Concepcion

Chile

Correspondence: Carlos Ossa-Cornejo. Casilla 447 Chillán, Chile. E-mail: cossa@ubiobio.cl

Abstract

Introduction. There are a few methods to study inquiry's abilities in Chile, despite its importance in science education. This study analyzes the psychometric properties of a Scientific Inquiry Scale in pedagogy students of two Chilean universities.

Method. The study uses an instrumental design with 325 students from 3 pedagogy majors. As a measurement instrument, it uses a scale based on one of the dimensions of Miranda's Tasks in Critical Thinking test. Descriptive statistics of reliability and dimensionality were used for the analysis.

Results. The results show three items under the mean, two within the mean and three slightly over it, with adequate ranges of asymmetry and kurtosis. Dimensions analysis accounts for a three-factor model with good fit indicators, related to the three ability foci observed in this ability: exploration with three items related, comprehension with two items and inference with three items. The reliability analysis is adequate but lower than expected (.61) and an adequate level of interjudge consistency, with an intraclass correlation coefficient of .62.

Discussion and Conclusion. Scientific Inquiry Scale is reliable and allows for analysis of the main elements of the ability to inquire, which is relevant for scientific reasoning. However, it is necessary to investigate the external validity.

Keywords: Scientific reasoning, psychometry, pedagogy, confirmatory factorial analysis

Resumen

Introducción. Existen escasos instrumentos para medir la habilidad de indagación en Chile, a pesar de su importancia para la Educación en Ciencias. Este estudio analiza las propiedades psicométricas de la Escala de Indagación Científica en estudiantes de Pedagogía de dos universidades chilenas.

Método. El estudio utiliza un diseño instrumental con 325 estudiantes de tres carreras de pedagogía. Como instrumento de medida se emplea una escala basada en una de las dimensiones del test Tareas de Pensamiento Crítico de Miranda. Para el análisis se utilizaron estadígrafos descriptivos, de confiabilidad y dimensionalidad.

Resultados. Los resultados muestran tres ítems bajo la media, dos dentro de la media y tres levemente por sobre ella, con adecuados rangos de asimetría y curtosis. El análisis de dimensiones da cuenta de un modelo de tres factores con buenos indicadores de ajuste, en relación con tres focos de habilidades observados en la habilidad medida, exploración con tres ítems relacionados, comprensión con dos ítems e inferencia con tres ítems. El análisis de confiabilidad es adecuado, pero bajo lo esperado (.61) y presenta un nivel de consistencia interjueces adecuado, con un coeficiente de correlación intraclase de .62.

Discusión. La escala de Indagación Científica es confiable y permite analizar los principales elementos de esta habilidad, que es relevante para el razonamiento científico. Sin embargo, se debe continuar investigando su validez externa.

Palabras Clave: Razonamiento científico, psicometría, pedagogía, análisis factorial confirmatorio

Introduction

Scientific reasoning is a set of cognitive abilities that permits analysis and evaluation of data that is obtained before a determined situation (Ding, 2014). It implies recognizing information that is obtained, distinguishing the elements that make it up, verifying the sources where it originates, and evaluating its coherence and relevance. For this, it is required to have an evaluative judgment about the observed situation and about the use of the cognitive abilities (Madariaga & Schaffernicht, 2013).

The education professionals should develop learning processes that promote these abilities, given that they permit to generate a responsible decision making, as well as to strengthen the abilities to fulfill the growing and complex requirements that appear in the scientific and professional fields (Gutierrez-Braojos, Salmerón-Vilchez, Martin-Romera & Salmerón, 2013; Juliá, 2006; Ku & Ho, 2010).

This can be achieved through the development of critical thinking abilities, which in the future education professionals should be a priority before the urgent measures of change that are required in the field of education, in order to promote a higher level of professionalism, development of policies that permit its promotion and its use in the school curriculum (Ávalos, 2007; Stapleton, 2011). This would serve, also, to improve the quality of learning, the academic performance, the autonomy and self-sufficiency of the students (Ku & Ho, 2010; Olivares, Saiz & Rivas, 2013; Tung & Chang, 2009).

The above is relevant due to the necessity to improve the quality of the initial teacher formation and elements of the professionalization that supports the effective development of knowledge by the teachers (Ávalos, 2007). This is seen reflected even more strongly in the performance of the teachers in the area of sciences and mathematics, which has been evidenced in the last international studies, where the Chilean students still have not reached the scientific and logic standards of the students of the OCDE countries (Organization for Economic Cooperation and Development, 2013) which could indicate that there is still a lot of work to be done in the training of teachers and in thus the achievement of critical thinking would be highly relevant (Ding, 2014).

Jointly with the previous, diverse proposals have been raised (Cofré et al., 2010; Cristobal & García, 2013; González, Martínez, Martínez, Cuevas & Muñoz, 2009; Harlen, 2013) in which the evaluation of the scientific inquiry competency has been made based on behavioral measures and in school science exams, existing scarce experiences that evaluate with quantitative instruments with adequate indicators of reliability and validity.

The scientific reasoning as an ability in the university formation

Scientific reasoning is an important ability for the formation of students in areas such as the sciences, technology, mathematics and engineering, for which a better formation should be delivered to facilitate the effective learning for the students, so that they can effectively apply their knowledge to the problems of reality (Bao et al., 2009). Nevertheless, today there is a crisis with respect to the level of value that is present in society and of the quality with which it is taught in the educational institutions (Benito, 2009). To improve this, the development of abstract thinking schemes is fundamental so that an application of knowledge is achieved in both the classroom as well as in daily life situations (Rodríguez, Mena & Rubio, 2010).

It has been proposed that the development of scientific thinking is related with advanced levels of thinking, in relation to the Piagetian formal operations stage, where abstract thinking predominates (Picquart, Guzmán & Sosa, 2010; Rodríguez et al., 2010). By the aforementioned, the teaching of science and scientific reasoning are probably related to the traditional methodologies like the master class (Ordoñez, 2014).

The development of activities that apply scientific abilities is necessary, for they strengthen its comprehension and usage (Akarsu, 2010). Some studies suggest that the development of application activities like research practice in the disciplinary fields constitute a situational factor that promotes the use of skills at the procedural and cognitive level (Benito, 2009; Ordoñez, 2014). Furthermore, when the comprehension of the techniques used is strengthened, it is possible to achieve a greater predisposition towards the scientific reasoning (Ordoñez, 2014) and even, in the achievement of greater levels of academic performance (Bao et al., 2009).

Scientific reasoning should permit the systematization in an ordered and logical manner of a set of analysis (Ding, 2014), using strategies, rules and plans that permit the devel-

opment of explanations about an observed phenomena (Lawson, 2004). What's more, it helps to judge the validity of knowledge that is generated, determining when to permit reaching a more accurate level of validity (Bao et al., 2009) which is also related to critical thinking (Saiz & Nieto, 2002).

Scientific inquiry

Inquiry is conceptualized as a cognitive process that implies making observations, formulating questions and examining sources of information to distinguish what is known and what is lacked knowing, also to use instruments to gather and analyze data, to propose answers, explanations and predictions (Garritz, 2010). It is a multifaceted activity that permits one to search for the construction and evaluation of one's own learning, implying the development of knowledge and understanding of scientific ideas (Camacho, Casilla & Finol, 2008).

It is understood as an important contribution to the education field to permit students to develop knowledge and understanding of scientific ideas in the management of their learning (Garritz, 2010). The inquiry model proposes that the teacher is an active agent in their own formation and assumes that the more that is known about the origins and consequences of their actions, the more possibilities can exist to make changes to it or to ground them in foundation (Miranda, Zambrano & Jélvez, 2010).

The capacity of inquiry would include different cognitive tasks oriented towards three main emphases, on one side, to identify and connect relevant information (extraction); on the other side, to analyze and reflect on the importance and validity of the information (comprehension and evaluation) (Miranda, 2003). On the other hand, the science teaching model based in inquiry signals five phases or stages for the development of this capability adding to these foci a previous process of focusing of the tasks and separating the application from the evaluation (Uzcátegui & Betancourt, 2013).

The use of the inquiry permits a greater comprehension of scientific concepts, as well as the development of scientific abilities, achieving a greater comprehension about the same nature of the science, in such achieving a comprehension of the development of scientific knowledge and its relationship to the society (González et al., 2009). The inquiry based learn-

ing promotes research and reflections on the pedagogical practice and its relationship with critical thinking (Ding, 2014).

The term scientific inquiry has been widely used in science education because of its closeness with search for valid data from the natural world, the generation of hypotheses, and the design of research. On the other side, it can be understood as an objective of learning, a teaching methodology or a pedagogical approach, that's to say a set of knowledge and beliefs that guide the teaching of science.

As a teaching model, it seems to be especially beneficial since it permits the construction of the class from the interests and questions of the same students inviting them to generate questions and hypothesis about daily life issues and their surroundings, to later support them in the elaboration and execution of the experimental designs oriented to prove their hypothesis and to provide answers to their questions (González et al., 2009).

Few instruments based in quantitative measurements of performance have been found for the evaluation of scientific inquiry, one of these is the New Practical Assessment Inventory NPTAI (Ferrés, Marbà, & Sanmartí, 2015), an adaptation of a test that evaluates scientific inquiry abilities based on closed-ended questions that assesses the achievement of students to complete the process of inquiry. The NPTAI allows for the analysis of 7 large categories that include the search for data, development of hypothesis, use of variables, selection of procedures, data processing, obtaining conclusions and meta-reflection.

While the instrument does integrate all the analysis levels of the inquiry process and many of those considered in the critical thinking, the search and selection processes, analysis and decision making, and meta-knowledge would be of different levels and functions for the theoretical models that critical thinking proposes (Miranda, 2003; Saiz & Nieto, 2002). In view of this, for the effects of the present work another instrument, the Inquiry Scale from the Tasks in Critical Thinking test (Miranda, 2003) has been selected for this study, as it permits an evaluation of the central abilities of inquiry (search and selection of data, analysis and inference), with adequate levels of reliability and greater parsimony.

Objetives and hypothesis

The objective of this study was to analyze psychometrics properties of scientific inquiry scale, adapted of Miranda's inquiry dimension of Critical thinking task test (2003) in Chilean's pedagogy students.

Method

Participants

The participants were 325 students from the second year of pedagogy in the mathematics pedagogy, natural sciences pedagogy, and history and geography pedagogy majors, from two traditional institutions of higher education in the region of Bio-Bio, Chile. The sample represent 55% of science and mathematic student's population in this area. Of these, 71% (231) were women, and 29% (94) men, presenting an age range from 17 to 38 years old (M=20.5, SD=2.54).

Instruments

For the study, a scale based in the inquiry dimension from the Tasks in Critical Thinking test from the Educational Testing Service of the United States was used, adapted and translated to Spanish by Miranda (2003). The instrument in general conceives critical thinking as a capability of selecting inferentially information, analyzing it to develop a hypothesis and to make decisions about how to communicate it. It has been used to evaluate the scientific thinking in biological and oceanographic areas of interest. The complete test consists of 14 questions, organized in three dimensions, which are inquiry, analysis and communication, structured in tasks that give an informative situation and questions (Miranda, 2003; Miranda et al., 2010).

The inquiry dimension consists of 8 items oriented to tasks such as identifying data of an act, searching systematically for relevant and valid data. This dimension presents an adequate level of reliability with a Cronbach's Alfa of .75 in the Chilean sample (Miranda, 2003). The items have an open-ended type of answer and its grading scores vary between 0 and 3 points, depending on the question, the maximum score being 20 points (Miranda et al., 2010); the scale is organized in three different tasks in a similar manner to the original test, understanding the first task as an analysis situation of graphic information with three questions. The

second task consists of a brief text and a graphic, considering two questions and finally the third task as a brief text with three questions.

Procedure

Before beginning the application, a revision by the ethics Committee of the Doctorate Program in Psychology of the Universidad de Concepcion, Chile of instrument, authorizations and informed consent was solicited. An application of the scale was completed without modification of the questions, being that it is had been applied in a Chilean population. For the application, a space during the class time was coordinated with the department heads and with previous authorization by the teacher in charge of the course; the instruments included a written informed consent that the students read before completing the test, and only if they voluntarily agreed did they proceed. Once received, the valid questionnaires were analyzed, that is to say those that had all the answers, or at least, only one missing answer per instrument. Afterwards, a database was generated with the results of the test through the statistical software SPSS version 20, where the statistical analysis was developed. Finally, a subsample of 105 subjects was extracted to complete an evaluation of the grading rubric of the scale based on the judgements of two independent evaluators.

Datal Analysis

To analyze the data, descriptive statistics were applied (central measures of tendency, dispersion and distribution), with the SPSS software V.20. Furthermore, an exploratory factorial analysis was completed using an extraction model of Factorization of the main axes and varimax rotation, in order to determine how many probable factors the scale has. Afterwards, a confirmatory factorial analysis was completed; in the evaluation of the global fit of the model the following statistics were considered as criteria: Chi squared (X^2), goodness of fit index (GFI), adjusted goodness of fit index (AGFI), comparative fit index (CFI), normed fit index (NFI), and the root mean square error of adjustment (RMSEA). Additionally, the Omega statistic was applied to measure the level of internal consistency, which is more adequate for data in which there could be differences between the data in the coefficient matrix, being more sensitive to those than Cronbach's Alfa (Campo, Villamil & Herazo, 2013); these were completed with the M Plus program, V.7.4, which permits the simultaneous estimation of the parameters for the proposed models like the modelization.

To estimate the consistency of the grading rubric, intraclass correlation coefficient was applied, with a randomized two-factor model and absolute agreement type, through the SPSS software V.20.0.

Results

The descriptive analysis of the inquiry scale items, as is shown in table 1, presents a range of answers in general, within the adequate parameters in terms of distribution with asymmetry and kurtosis values between 2 and -2, which is the expected for a normal distribution according to Bollen and Long (cited in Nuñez-Alonso, Martín-Albo & Navarro, 2007). It is observed that items 6 and 8 present elevated kurtosis, which could indicate that the results would be largely concentrated near the mean. See Table 1.

Table 1.

Descriptive analysis of the items

	N	Minimum	Maximum	Mean	Std. Deviation	Asymmetry	Kurtosis
Item1	325	0	3	1.39	.834	.360	-.401
Item2	325	0	3	1.33	.715	.642	.308
Item3	325	0	2	1.22	.600	-.133	-.474
Item4	325	0	3	1.58	.743	.252	-.437
Item5	325	0	2	1.38	.682	-.651	-.687
Item6	325	0	2	1.20	.605	.884	1.546
Item7	325	0	3	1.22	.630	.455	.570
Item8	325	0	2	.91	.446	-.404	1.728

Related to the average of each item, it is observed that in questions 1, 2 and 7 the value is lower than the expected arithmetic mean (1.5 points), the same for item 8 (1 point); while in question 4 the value is approximately the arithmetic mean; questions 3, 5 and 6 show a higher average than the expected arithmetic mean (1 point). In the analysis of model ‘s dimensions, the exploratory factorial analysis (EFA) showed a KMO of .63 which is considered only adequate, while the Bartlett test of sphericity was significant (p=.000). Three factors emerged that explain 32% of the variance.

For the confirmatory factorial analysis (CFA), three models were tested, with different amounts of factors, as is observed in table 3, the model of three factors is that which presents the best fit indexes (values CFI, GFI, NFI and AGFI optimal over 0.95; RMSEA optimal under 0.05). The previous would be in agreement with the proposed structure by the EFA as well as with the organization of the test in three areas.

Table 2.

CFA values of inquiry scale

Models	X^2 (p)	CFI	GFI	NFI	AGFI	RMSEA
One Factor	11.031 (.68)	1.00	.955	.957	.980	.00
Two Factors	24.585 (.175)	.975	.929	.904	.965	.038
Three Factors	13.361 (.77)	1.00	.965	.965	.982	.00

It can be observed in the model diagram below (figure 1), that the three factors present adequate correlations between them, only that factor 1, which is named “exploration” (In 1 in figure), has a significant relationship but of low intensity with the factor 3 that corresponds to “inference” (In 3 in figure), but not with factor 2, that is named “comprehension” (In 2 in figure). A very strong correlation is observed between factors 2 and 3, which could explain the analytic nature of these tasks.

The model presented furthermore demonstrates that is it adequate to establish the scale with three types of tasks, in function of the abilities that are present in the inquiry. Items 1, 2 and 3 would have a significant relationship and of middle to high with the exploration factor, while items 4 and 5 present a significant and middle correlation with the comprehension factor. Finally it is observed that items 6, 7 and 8 present a significant relationship with the inference factor; even though only 7 and 8 show high correlation, in exchange item 6 is observed with a low intensity relationship. See Figure 1.

Finally, the omega statistic, shows a medium reliability level of the inquiry scale, lower that than expected but still positive (α=.613). While the intraclass correlation analysis obtained, based on the evaluation of the two judges in the subsample of students (N=103,

M1=6.9, SD=3.10; M2=9.01, SD=3.11), was .618, which accounts for an adequate consistency between the judges.

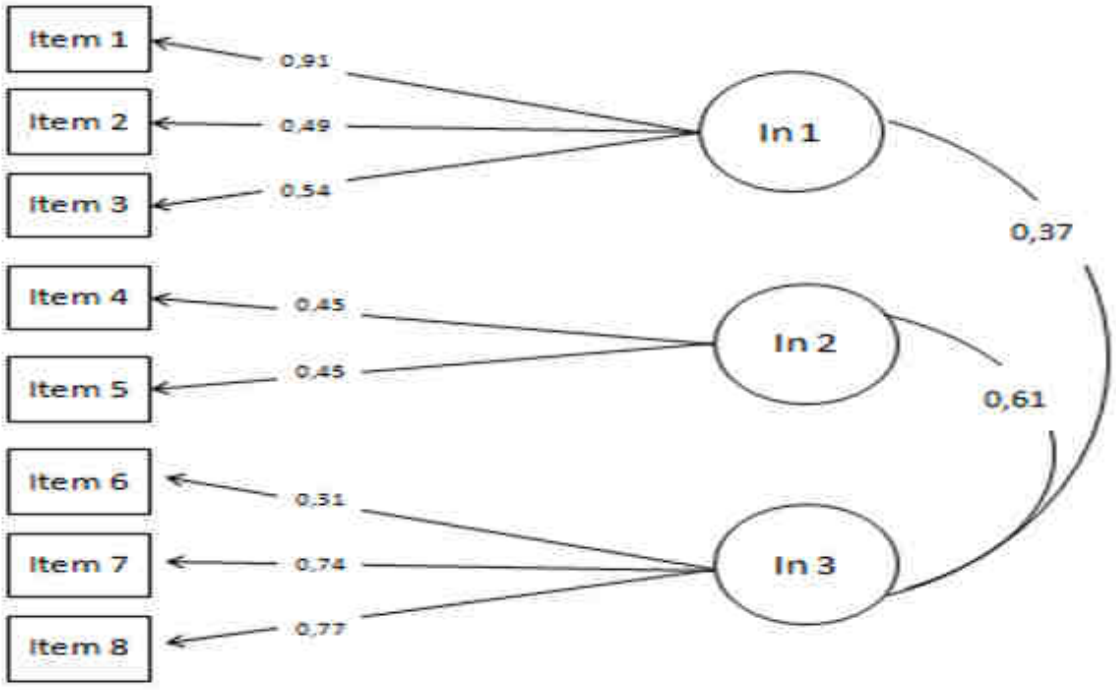

Figure 1. Three-factor model of the inquiry scale

Discussion and Conclusions

The development of the scientific inquiry is a relevant topic in the current formative strategies in the science teaching (Bao et al., 2009; Benito, 2010; González et al., 2009; Rodríguez et al., 2010). It is possible to find currently a larger emphasis in the study of programs and didactic experiences for the development of the inquiry competency, on the one hand, and in the determination of proficiency profiles required to carry out the inquiry on the other (Cristóbal & García, 2013; Harlen, 2013), abounding instruments of observational and qualitative registers, and with less interest in developing instruments that allow for the establishment of a performance measurement with a higher reliability. This point is relevant since having measurement instruments permits the replication of the evaluation of programs and

experiences with a more reliable base, generating greater possibilities of comparison. Moreover, it decreases the variability of the perceptions in the student performance.

The measurement instrument examined in this study presents adequate indicators of reliability, while it is found lower than expected in comparison with other studies (Miranda, 2003; Miranda et al., 2010) and the conventional minimum of .70 (Barraza, 2007). The inter-judge consistency is adequate but also somewhat low. According to what the previous author mentioned, the indicators for the instrument would be within the acceptable level.

The dimension analysis proposes that it is not possible to consider this scale as only one factor, being that its analysis shows that the indicators don't fit adequately to the model. This is contrary to what is mentioned in the validation study of the instrument (Miranda, 2003), that integrates the items in one unique dimension. Nevertheless, from the results of this study, it was observed that it is more adequate to consider three dimensions, achieving indicators with a good fit model and considering furthermore that the items are oriented to different aspects of the inquiry task.

It is worth mentioning that this proposed tripartite structure is consistent with the approach developed by the ECBI model (Uzcátegui & Betancourt, 2013), focusing on the three central processes, even when the evaluation phase proposed in the model is not considered. It is also important to note the adequate existing relationship between the items and the proposed factors, even though it emphasizes the fact that between them there is an adequate relationship between the Exploration and Inference dimensions, and between the Comprehension and Inference but not between Exploration and Comprehension. The above would indicate that in the type of tasks that are integrated in the instrument, the abilities to explore and to comprehend are both fundamental for the inference of information, however, the exploration in detail would not be necessarily necessary to comprehend the information.

It is observed as limitations of the study the instrumental character of the same, oriented to analyze aspects of internal validity of the instrument rather than to analyze the data generated. On the other side, it is relevant to note, as a deficiency, the level of reliability found in the items, as well as in the evaluation of the instrument and the reviewers of the instrument, which does not allow to establish a categorical assessment of the scale as a useful and accu-

Electronic Journal of Research in Educational Psychology, 15(3), 624-640. ISSN: 1696-2095. 2017. no. 43
http://dx.doi.org/10.14204/ejrep.43.17027

rate instrument of measurement. The existence of difficulties in writing the items is hypothesized which generates difficulty in the understanding of the tasks to be carried out.

Finally, as projections of the study its worth mentioning, in relation to the above, the need to improve the writing of the items and to improve the level of reliability and interjudge consistency. In addition, it is necessary to advance in external validation through tests of concurrent and discriminant validity with other instruments that evaluate attention abilities, perceptive discrimination, inductive and deductive reasoning and inference.

Acknowledgments

This research was developed with support of CONICYT FONDECYT Chilean fund nº 1161502 “Modelo explicativo de la permanencia y el abandono de los estudios universitarios, basado en procesos cognitivo motivacionales”, and CONICYT Chile (CONICYT-PCHA/Doctorado Nacional/2015-21150389).

References

Arenas, M. (2000). Diseño instrumental de un modelo factorial para predecir el rendimiento y éxito profesional del comunicador visual. *Investigación Universitaria Multidisciplinaria: Revista de Investigación de la Universidad Simón Bolívar*, *2,* 30-36. Retrieved from: https://dialnet.unirioja.es/descarga/articulo/3683913.pdf

Akarsu, B. (2010). Turkish pre-service teacher’s perspectives of demonstrations and hands-on activities in science classrooms. *Latin-American Journal of Physics Education, 4*(3), 506–510. Retrieved from: http://www.lajpe.org/sep10/431_Bayram_Akarsu.pdf

Ávalos, B. (2007). El desarrollo profesional continuo de los docentes: lo que nos dice la experiencia internacional y de la región latinoamericana. *Revista Pensamiento Educativo, 41*(2), 77–99. Retrieved from: http://pensamientoeducativo.uc.cl/index.php/pel/article/view/417/854

Bao, L., Cai, T., Koenig, K., Fang, K., Han, J., Wang, J., … & Wu, N. (2009). Learning and scientific reasoning. *Science*, *323*(5914), 586–587. http://doi.org/10.1126/science.1167740

Benito, M. (2009). Debates en torno a la enseñanza de las ciencias. *Perfiles Educativos*,

XXXI(123), 27–43. Retrieved from: http://www.scielo.org.mx/pdf/peredu/v31n123/v31n123a3.pdf

Camacho, H., Casilla, D. & Finol , M. (2008). La indagación: una estrategia innovadora para el aprendizaje de procesos de investigación. *Laurus*, *14*(26), 284–306. Retrieved from: http://www.redalyc.org/pdf/761/76111491014.pdf

Cofré, H., Camacho, J., Galaz, A., Jiménez, J., Santibánez, D. & Vergara, C. (2010). La educación científica en Chile: debilidades de la enseñanza y futuros desafios de la educacion de profesores de ciencia. *Estudios Pedagógicos*, *XXXVI*(2), 279–293. http://doi.org/10.4067/S0718-07052010000200016

Cristobal, C. & García, H. (2013). La indagación científica para la enseñanza de las ciencias. *Horizonte de La Ciencia*, *3*(5), 99–104. Retrieved from: http://dialnet.unirioja.es/descarga/articulo/5420523.pdf

Ding, L. (2014). Verification of causal influences of reasoning skills and epistemology on physics conceptual learning. *Physical Review Physics Education Research*, *10*(2), 1–5. http://doi.org/10.1103/PhysRevSTPER.10.023101

Ferrés, C., Marbà, A. & Sanmartí, N. (2015). Trabajos de indagación de los alumnos: instrumentos de evaluación e identificación de dificultades. *Revista Eureka Sobre Enseñanza y Divulgación de Las Ciencias*, *12*(1), 22–37. http://doi.org/10498/16922

Garritz, A. (2010). Indagación: las habilidades para desarrollarla y promover el aprendizaje. *Educacion Quimica*, *21*(2), 190–197. Retrieved from: http://andoni.garritz.com/documentos/2013/04_editVol21-2Indagacion2010.pdf

González, C., Martínez, M. T., Martínez, C., Cuevas, K. & Muñoz, L. (2009). La educacion cientifica como apoyo a la movilidad social: Desafios en torno al rol del profesor secundario en la implementacion de la indagación científica como enfoque pedagógico. *Estudios Pedagógicos*, *35*(1), 63–78. http://doi.org/10.4067/S0718-07052009000100004

Gutierrez-Braojos, C., Salmerón-Vilchez, P., Martin-Romera, A. & Salmerón, H. (2013). Efectos directos e indirectos entre estilos de pensamiento, estrategias metacognitivas y creatividad en estudiantes universitarios. *Anales de Psicologia*, *29*(1), 159–170. http://doi.org/10.6018/analesps.29.1.124651

Harlen, W. (2013). *Evaluación y educación en ciencias basada en la indagación: Aspectos de la política y la práctica*. Trieste: Global Network of Science Academies (IAP) Science Education Programme (SEP). Retrieved from: http://www.fondation-lamap.org/sites/default/files/upload/media/IBSE assessment guide _ spanish.pdf

Juliá, M. T. (2006). Competencias profesionales del psicólogo educacional: una tarea

asociativa. *Revista de Psicología, XV*(2), 115–130. Retrieved from: http://www.redalyc.org/pdf/264/26415207.pdf

Ku, K. Y. L. & Ho, I. T. (2010). Metacognitive strategies that enhance critical thinking. *Metacognition and Learning, 5*(3), 251–267. http://doi.org/10.1007/s11409-010-9060-6

Lawson, A. E. (2004). The nature and development of scientific reasoning: A synthetic view. *International Journal of Science and Mathematics Education, 2*(3), 307–338. http://doi.org/10.1007/s10763-004-3224-2

Madariaga, P. & Schaffernicht, M. (2013). Uso de objetos de aprendizaje para el desarrollo del pensamiento crítico. *Revista de Ciencias Sociales (Ve), XIX*(3), 472–484. Retrieved from: http://www.redalyc.org/pdf/280/28028572010.pdf

Miranda, C. (2003). El pensamiento crítico en docentes de educación general básica en Chile: un estudio de impacto. *Estudios Pedagógicos*, 29, 39–54. http://dx.doi.org/10.4067/S0718-07052003000100003

Miranda, C., Zambrano, F. & Jélvez, M. (2010). ¿Incide la formación inicial en el desarrollo del pensamiento crítico de los estudiantes de pedagogía? Insumos desde un estudio de caso para un debate en curso. *Boletín de Investigación Educacional, 24*(1), 79–89.

Montero, I. & León, O. (2005). Sistema de clasificación del método en los informes de investigación en Psicología. *International Journal of Clinical and Health Psychology, 5*(1), 115–127. Retrieved from: http://www.aepc.es/ijchp/SCMIIP.pdf

Olivares, S., Saiz, C. & Rivas, S. (2013). Encouragement for thinking critically. *Electronic Journal of Research in Educational Psychology, 11*(30), 367–394. http://doi.org/10.14204/ejrep.30.12168

Ordoñez, O. (2014). Replicar para comprender: prácticas investigativas para promover el razonamiento científico en estudiantes de psicología. *Pensamiento Psicológico, 12*(2), 7–24. http://doi.org/10.11144/Javerianacali.PPSI12-2.rcpi

Organisation for Economic Cooperation and Development. (2013). *Assessment of Higher Education Learning Outcomes (AHELO) Feasibility Study report - Volume 3 further insights*. Paris: Organisation for Economic Cooperation and Development. Retrieved from:http://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume3.pdf

Picquart, M., Guzmán, O. & Sosa, R. (2010). Razonamiento científico e ideas previas en alumnos de ciencias básicas de la UAM- Iztapalapa. *Latin-American Journal of Physics Education, 4*(1), 1056–1064. Retrieved from:

https://dialnet.unirioja.es/descarga/articulo/3700627.pdf

Rodríguez, M., Mena, D. & Rubio, C. (2010). Razonamiento científico y conocimientos conceptuales de mecánica: un diagnóstico de alumnos de primer ingreso a licenciaturas en ingeniería. *Formación Universitaria*, *3*(5), 37–46. http://doi.org/10.4067/S0718-50062010000500006

Saiz, C. & Nieto, A. M. (2002). Pensamiento crítico: capacidades y desarrollo. In C. Saiz (Ed.), *Pensamiento crítico: conceptos básicos y actividades prácticas* (pp. 15–19). Madrid: Piramide.

Stapleton, P. (2011). A survey of attitudes towards critical thinking among Hong Kong secondary school teachers: Implications for policy change. *Thinking Skills and Creativity*, *6*(1), 14–23. http://doi.org/10.1016/j.tsc.2010.11.002

Tung, C. A. & Chang, S.Y. (2009). Developing Critical Thinking through Literature Reading. *Feng Chia Journal of Humanities and Social Sciences*, 19, 287–317. Retrieved from: http://www.cocd.fcu.edu.tw/wSite/publicfile/Attachment/f1262069682958.pdf

Uzcátegui, Y. & Betancourt, C. (2013). La metodología indagatoria en la enseñanza de las ciencias: una revisión de su creciente implementación a nivel de Educación Básica y Media. *Revista de Investigación, 37*(78), 109–127. Retrieved from: http://www.scielo.org.ve/pdf/ri/v37n78/art06.pdf