

Número 43

Enero-Junio 2015

Presentación

Eloísa Alcocer Vázquez
Universidad Autónoma de Yucatán 4

Artículos

How can they be gifted if they don't speak English? Teacher Perceptions of the Talents of Second Graders in a Dual Language Program

Liz Hollingworth y Sarah Hale Keuseman
The University of Iowa 6

Caracterización de directores de escuelas primarias del sur de Sonora

Ana Patricia Luna Gómez, Sonia Beatriz Echeverría Castro y José Javier Vales García
Instituto Tecnológico de Sonora 19

Influencia del estrés académico en el rendimiento escolar de un grupo de estudiantes universitarios

Rodrigo Arturo Domínguez Castillo, Gladys Julieta Guerrero Walker y José Gabriel Domínguez Castillo
Universidad Autónoma de Yucatán 31

Relación entre aprendizaje de conceptos y aprendizaje de procedimientos en economía de la empresa agraria

Luis Alberto Araujo Andrade y Eraclio del Jesús Cruz Pacheco
Universidad Autónoma de Yucatán 41

Propiedades psicométricas de un instrumento para medir la dinámica familiar en niños

Esthela Jacqueline Madrid López, Ángel Alberto Valdés Cuervo y María Dolores Valadez Sierra

Instituto Tecnológico de Sonora, Universidad de Guadalajara

54

Reflexiones

Challenges Measuring Intelligence in Mexico: Lessons Learned from the Standardization of the Wechsler Scales

Pedro Antonio Sánchez Escobedo

Universidad Autónoma de Yucatán

65

Sobre la técnica de puntajes de propensión (Propensity Score Matching) y sus usos en la investigación en educación

Claudia Patricia Ovalle Ramírez

Universidad de Antioquia

80



<http://www.educacionyciencia.org>
educación.ciencia@correo.uady.mx

Equipo editorial

Dra. Eloísa Alcocer Vázquez
Directora

Dr. Pedro Sánchez Escobedo
Editor científico

Mtro. Israel Alberto Cisneros Concha
Editor de producción

MINE. Hugo Salvador Flores Castro
Editor de innovación

Consejo editorial

Dra. Graciela Cordero Arroyo
Universidad Autónoma de Baja California

Dra. Silvia Joaquina Pech Campos
Universidad de Castilla-La Mancha

Dra. Liz Hollingworth
University of Iowa

Dr. Pascal Lafont
Universidad Paris-Est Créteil Val de Marne

Dra. Mariela Sonia Jiménez Vázquez
Universidad Autónoma de Tlaxcala

Mtra. Hilda Marisela Partido Calva
Universidad Veracruzana

Dra. Leticia Pons Bonals
Universidad Autónoma de Chiapas

Dra. Edith Cisneros Cohernour
Universidad Autónoma de Yucatán

Dr. Pedro José Canto Herrera
Universidad Autónoma de Yucatán

Dra. Zoia Bozu
Universidad de Barcelona

Dr. Rubén Comas Forgas
Universidad de las Islas Baleares

Dr. Javier José Vales García
Instituto Tecnológico de Sonora

Mtra. Elizabeth Narváez Cardona
Universidad Autónoma de Occidente

Dra. María Isabel Ocampo Tallavas
Universidad Autónoma "Benito Juárez" de Oaxaca

Presentación

Educación y Ciencia cumple veinticinco años de estar difundiendo investigaciones originales en el área educativa. En el transcurso de este tiempo, *Educación y Ciencia* se ha renovado para estar a la vanguardia del conocimiento científico y generar un impacto positivo en el quehacer educativo. El número 43 continúa con la larga tradición de la labor de la revista y reafirma el compromiso hecho desde 1990: ser un canal abierto de comunicación para la comunidad académica y la sociedad en torno a la investigación educativa.

El número 43 abre con el artículo de Liz Hollingworth y Sara Hale Keuseman sobre la situación de los *English Learners* en los Estados Unidos, la cual implica un reto para el actual contexto de diversidad cultural de ese país. Como las mismas autoras expresan, este artículo se enmarca en la exigencia de justicia social y equidad que sale a la luz cuando se exploran las causas de la falta de representación de minorías lingüísticas en los programas y distinciones escolares para estudiantes talentosos y superdotados. Los *English Learners* tienen habilidades especiales desarrolladas debido a particular circunstancia social, tales como ser bilingües o políglotas; sin embargo, las autoras apuntan que éstas son muchas veces subestimadas en las mediciones estandarizadas para la búsqueda de estudiantes sobresalientes.

Ana Patricia Luna Gómez, Sonia Echeverría Castro y José Vales García hacen un análisis sobre el perfil de los directores de escuelas primarias en el sur de Sonora con el objetivo de observar cuáles son las características que se relacionan con un alto o bajo nivel de desempeño académico de la escuela. En el director recae la responsabilidad de gestión, organización, planificación y evaluación. Por lo tanto, la función directiva se convierte en un elemento sustancial para el día a día de cada escuela y la educación de calidad. Los autores reflexionan sobre la necesidad de una capacitación formal para la administración escolar.

En los dos artículos siguientes, los estudios se adentran en la población universitaria. Primero, Rodrigo Domínguez Castillo, Julieta Guerrero Walker y Gabriel Domínguez Castillo evalúan el concepto de estrés y su relación con el rendimiento académico. La justificación del estudio se enmarca en un cambio profundo y actual en la conducta diaria de niños, jóvenes y adultos debido a factores económicos y sociales, entre otros. De esta manera, el estrés es un tema inmediato a nuestro contexto social que afecta nuestras percepciones e interacciones sociales. Posteriormente, Luis Araujo Andrade y Eraclio Cruz Pacheco exploran la relación que existe entre el aprendizaje de conceptos y el aprendizaje de procedimientos de acuerdo a la taxonomía de Bloom. Los investigadores presentan el valor de ambos conocimientos y su correlación para el desarrollo profesional óptimo de los estudiantes.

Por su parte, Esthela Madrid López, Angel Valdés Cuervo y María Dolores Valadez Sierra analizan la confiabilidad y validez de un instrumento para medir la dinámica familiar en niños de escuelas primarias. Los resultados de este estudio revelan la importancia del rol de los padres de familia en el desarrollo sano de los niños durante su infancia y en el ambiente escolar.

A continuación, Pedro Sánchez Escobedo inaugura la sección de Reflexiones de *Educación y Ciencia* que es un espacio para el análisis riguroso y crítico de teorías, metodologías y prácticas del área de educación. En esta ocasión, el artículo de Sánchez Escobedo examina a profundidad la experiencia mexicana en la adaptación de las escalas de Wechsler usadas para la medición de la inteligencia. El artículo explora las lecciones aprendidas, las ventajas y desventajas de la adaptación y estandarización de estas tres escalas que tiene su origen en un

contexto diferente al mexicano. Por su parte, Claudia Ovalle Ramírez analiza los beneficios del uso de la técnica estadística Propensity Score Matching para la investigación educativa.

Como se puede observar, el número 43 tiene un aporte significativo para entender y evaluar el quehacer educativo. De esta manera, los invitamos a celebrar los veinticinco años de *Educación y Ciencia* y su compromiso por generar y comunicar el conocimiento científico a través de la lectura, el análisis y los comentarios a los contenidos.

Eloísa Alcocer Vázquez
Universidad Autónoma de Yucatán



¿Cómo pueden ser altamente sobresalientes si no hablan inglés? Percepciones de los maestros de alumnos con talento de segundo grado en un programa bilingüe

How Can They Be Gifted If They Don't Speak English? Teacher Perceptions of the Talents of Second Graders in a Dual Language Program

Liz Hollingworth y Sarah Hale Keuseman
The University of Iowa

Información adicional sobre este manuscrito escribir a:

Liz Hollingworth, liz-hollingworth@uiowa.edu, Sara Hale Keuseman, sarah-hale@uiowa.edu

Cómo citar este artículo:

Hollingworth, L., & Keuseman S. H. (2015). How Can They Be Gifted If They Don't Speak English? Teacher Perceptions of the Talents of Second Graders in a Dual Language Program. *Educación y ciencia*, 4(43), 6-18.

Resumen

Existe una necesidad de evaluaciones que puedan proporcionar a los educadores valiosa información para lograr intervenciones educativas que ayuden a los estudiantes a alcanzar su potencial académico, en particular a los estudiantes de grupos lingüísticos minoritarios. En este estudio se detalla la relación entre las calificaciones asignadas por los profesores al desempeño de los estudiantes y los resultados en las pruebas estandarizadas de rendimiento escolar y habilidades académicas. Las preguntas de investigación fueron: ¿en qué medida coinciden las calificaciones de los estudiantes asignadas por los maestros y el rendimiento en las pruebas estandarizadas de logros y capacidades académicas?; ¿qué tipo de datos son necesarios para tomar decisiones acerca de la distinción como talentosos y sobresalientes a estudiantes en aprendizaje del inglés?; y ¿qué talentos son pasados por alto o subestimados utilizando los datos que están disponibles para los maestros? Este trabajo se basa en la creciente literatura que aborda una perspectiva social más justa para la identificación y la educación de los niños sobresalientes y talentosos en aprendizaje del inglés.

Palabras clave: educación bilingüe, evaluación, estudiantes sobresalientes

Abstract

There is a need for assessments that can provide educators with valuable information to allow for educational interventions for students to achieve their academic potential, particularly students from linguistic minority groups. This study details the relationship among teacher ratings of student academic achievement and standardized test achievement and ability measures. The research questions were: to what extent do the teacher ratings of students and the performance on standardized tests of achievement and ability match; what kinds of data are needed for teachers to make decisions about gifted programming for English Learners; and what talents are overlooked/undervalued using data that are available to teachers? This work builds on the growing literature base informing a more socially just policy for the identification of and education of gifted and talented children learning English.

Keywords: bilingual, assessment, gifted, talented, English Language Learners

Introduction

English Learners (ELs) are students “whose English proficiency has not yet developed to a point where they can profit fully from English instruction” (García, Jensen, & Scribner, 2009, p. 10). Today, more than 10% of U.S. students are EL, an increase of 150% from 20 years ago, and ELs are the fastest-growing subgroup in the nation (Spellings, 2005; U.S. Department of Education, 2012). By 2025, researchers project ELs will comprise a quarter of all students (Spellings, 2005). Despite the increase in the population, ELs are disproportionately underrepresented in gifted and talented programs across the country (Callahan, 2005; Ford, 1998).

One way American schools determine whether a student qualifies for gifted and talented programming is through the use of standardized tests. Specifically, ability tests are widely used in American schools and provide important information that cannot be derived from achievement tests, teacher observations, or student self-ratings of learning styles. Because of this, ability tests are used to help educators make better decisions about students’ academic programming. For example, students who show unusual discrepancies between their scores on achievement and ability tests may be identified for evaluation of a learning disability. Conversely, students with unexpectedly high ability in one or more domains may be provided access to educational programming and resources reserved for gifted and talented students. Commonly used screening tests for these purposes in K-12 schools include the Otis-Lennon School-Ability tests (Otis & Lennon, 2003), Inview (CTB/McGraw-Hill, 2000), the Progressive Matrices Test (Raven, Court, & Raven, 1983), the Naglieri Nonverbal Ability test (Naglieri, 1997) and the Cognitive Abilities Test (CogAT; Lohman & Hagen, 2001). A national survey by MMS Education (2007) found that the CogAT was by far the most widely used of these tests. Typically, the directions and/or items on ability tests, which often have novel item types, are difficult to understand if the student is not fluent in English.

Many have assumed that nonverbal ability tests would level the playing field for non-native English speakers. However, Lohman, Korb, and Lakin (2008) found that EL students scored 8 to 10 points lower than native English speakers on three nonverbal tests. Thus, for this project an adapted form of the CogAT Form 7 with Spanish instructions was used with selected students. Teachers identified which students would receive the Spanish directions based on home language and teacher perception of ability in English and Spanish.

One of the most important features of CogAT is the profile of each student’s abilities to reason with verbal, quantitative, and figural/spatial symbols. Profiles with extreme score discrepancies are particularly common for both high- and low-ability students (Lohman, Gambrell, & Lakin, 2008). For low-ability students, score profiles typically show a significant cognitive strength whereas for high-ability students profiles often show a relative weakness. In both cases, students learn more successfully when instruction is adapted better to capitalize on their strengths (Snow, 1977).

Policy Context

In 1965, the United States Congress passed the largest federal intervention in education to that date, the Elementary and Secondary Education Act (ESEA). That original version of ESEA was intended to provide compensatory programs for students from lower socioeconomic levels. In the 1965 ESEA, Title III addressed innovative programs in education. Three years after the

original passing of the bill, ESEA was amended to include Title VII, the Bilingual Education Act with outcomes to better educate children who did not speak English as a first language “to encourage the establishment and operation, where appropriate, of educational programs that use Bilingual educational practices, techniques, and methods” and “to provide financial assistance to local education agencies, and to state education agencies for certain purposes” (Iowa Department of Education, 2007, p. 3).

Throughout the 1970s, 80s, and 90s the federal government’s goals for educating English Learners (ELs) changed. By 1973 the goals of Title III had developed into providing money to short-term programs (often no longer than three years) to school districts (Norman & Balyeat, 1973). A series of court cases, including *Lau v. Nichols*, 1974, *Castenada v. Pickard*, 1981, and *Plyler v. Doe*, 1982 emphasized the legal requirements to educating EL students, as well as students whose immigration status is undocumented. By the time the 1994 reauthorization of the ESEA passed, funding for EL programs was provided through a series of competitive grants through Title VII (Boyle, Taylor, Hurlburt, & Soga, 2010a). The focus of EL education had also moved from bilingual education to a focus on English language acquisition (Ramsey & O’Day, 2010).

A major change to the education of ELs came with the 2001 reauthorization of ESEA as the No Child Left Behind Act (NCLB). The act outlined a new paradigm for education, centered on accountability, “Now, in addition to detailed grade level expectations for student achievement in reading and math, states were also required to set academic achievement standards with cut-scores for basic, proficient and advanced performance” (Hollingworth, 2008). Education for ELs moved from Title VII to Title III, and funding changed from a competitive grant structure to formula-based grants allocated to states, depending on the population of EL students in that state. States were to issue grants to districts or consortia of districts to implement the programs for ELs. With this funding, grantees were required use the monies to enhance language instruction educational programs to improve English Language proficiency and academic achievement and to provide professional development to teachers surrounding EL instruction (Boyle, Taylor, Hurlburt, & Soga, 2010b). In return for these monies, schools needed to demonstrate that they had a system for holding educators accountable for student learning.

From a federal policy perspective, both Title I of the 1994 Improving America’s Schools Act and NCLB require schools to test all students to establish annual goals for accountability purposes, even those who are limited in their ability to express themselves in English. However, there is limited research on the value of those student test scores to make educational decisions that might benefit this specific population. Lindholm-Leary and Borsato (2006) found “while there has been an increased emphasis on the inclusion of ELLs [English Language Learners] in high-stakes tests, research on assessment that is pertinent to this population is scarce” (p. 196). Many school systems spend significant time and effort on making sure that EL students who are falling behind are able to close the gap to meet assessment standards. However, research and practices surrounding EL students who meet the criteria for gifted and talented is limited.

Research on Gifted Bilingual Students

The construct of giftedness from a policy perspective is typically defined in terms of academic, school subject-specific talents, as opposed to a child’s fluencies in multiple languages. Valdés and Figueroa (1994) write, “the unique American tragedy of bilinguals has been that over the last century both test makers and testers have generally ignored the psychological robustness of bilingualism. The result has been a waste of human potential” (p. 87). Vald and Valdes (2003)

outline the debates in gifted education about the inclusion of young bilingual interpreters in the construct of “giftedness”. The authors argue “formal definitions of giftedness are particularly important because they become part of official policies and guidelines that are later used to inform identification and programming practices in schools” (p.8). Indeed, current policies for identification have resulted in the low representation of linguistic minority students in gifted programs. Bialystok (2007) contends bilingualism is advantageous because “bilinguals develop executive control earlier and maintain their ability to control those functions longer than monolinguals” (p. 220). Gonzalez (2006) found that bilingual, low SES children had the cognitive potential to outperform monolingual, middle-to-high SES counterparts.

However, EL and bilingual children have long been under-represented in programs for gifted and talented (G&T) students in American public schools (Barkan & Bernal, 1991; Callahan, 2005; Ford, 1998). Esquierdo and Arreguin-Anderson (2012) go so far as to identify these students as “the invisible gifted and talented bilingual students”. They argue that two factors in this underrepresentation include the definition of giftedness that many school systems use and teachers’ level of preparation in gifted education. Castellano and Diaz (2002) identify additional reasons for the underrepresentation of Hispanic students in gifted education programs, including low expectations from educators, non-responsive curriculum, and inadequate identification and assessment tools. This lack of adequate assessment tools is a significant barrier to accurate identification (Solano-Flores & Trumbull, 2003): “Existing approaches to testing ELLs do not ensure equitable and valid outcomes because current research and practice assessment paradigms overlook the complex nature of language, including its interrelationship with culture” (Solano-Flores & Trumbull, p. 3).

Moreover, there is a lack of use of students’ first language in assessing academic abilities across content areas (Gonzalez, 2012). The lack of first language assessment is especially concerning as primary language is connected to “personal and cultural identity, to the expression of their learning styles and temperament, self-concept, and self-esteem, all factors very much connected to the cultural expression of their cognitive and academic skills” (p. 293). Gifted and talented assessment tools essentially use language to identify students for gifted and talented programs and it is critically important to recognize the difference between the students’ inability to express knowledge and ability as opposed to the actual lack of knowledge or ability (Fultz, Lara-Alecio, Irby, & Tong, 2013). This difference amplifies the possibility of measurement error and the risk of misidentifying students for entry to gifted and talented educational programs (Solano-Flores & Trumbull, 2003).

These programs range from pull-out classes, to special schools, to enrichment programs open to most any student in the school who displays interest a topic and the ability to pursue it. Collectively, in-school talent identification and development programs are an important resource for all students, especially for poor children whose parents cannot afford or who are reluctant to participate in special tutoring, field trips, summer programs, and other enrichment and acceleration opportunities. Although the criteria for admission to such programs vary widely across school districts, many districts administer an ability test in first or second grade either to those students who are nominated for the program by a classroom teacher or to all students in the identified grade level.

Research Questions

We wondered if the teachers and the existing achievement tests could identify gifted and talented children in a school district that has embraced the values of bilingual education for all elementary children. The research questions guiding this study are:

1. To what extent do the teacher ratings of students and the performance on standardized tests of achievement and ability match?
2. What kinds of data are needed for teachers to make decisions about gifted and talented placement for children who are ELs?
3. What talents are overlooked or undervalued using the data that are available to the teachers?

Methods

The case study is designed to gain an in-depth understanding of both the situation and the meanings for the subjects in order to develop general theoretical statements about the process under observation (Merriam, 1998). Yin (2003) also recommends relying on theory to drive the analysis of case study data.

Primary Research Site

The Community School District (CSD) is located in a Midwestern town with a population of 3,531. The community is the first minority-majority town in the state, with 52% of the population identifying as Hispanic on the 2010 U.S. Census. There are 1,168 K-12 students in the district, which spends \$4,231 per pupil. Table 1 identifies the participant characteristics. Five languages are spoken by PK-12th grade students in CSD: English, Spanish, Lao, Vietnamese, and Nepali. More than one-third of the students list a language other than English spoken in their home.

Table 1. Participants Characteristics

Participants	N	%
Students (Grade 2)		75
Hispanic	37	49
Home Language English	47	63
Home Language Other	29	39
Special Education	7	93
Gifted and Talented	0	0
Free and Reduced Lunch	44	59
Receiving Spanish directions	20	27
English Learners	25	33.
Teachers		11

In October 2010, 13% of the district 2nd graders qualified for special education services. None were identified as talented and gifted, as the program starts in 3rd grade. Table 2 shows the percent's proficient in reading and math of the fourth grade class, as reported by the state for NCLB purposes.

Table 2. Iowa Tests of Basic Skills Percent Proficiency 4th Grade- Biennium 2007-2009

	Reading		Math	
	District	State	District	State
ELL	52.27	54.72	50.00	58.40
African American	N/A	57.66	N/A	56.26
White	78.87	82.13	78.87	83.55
Hispanic	52.78	63.87	52.78	64.98
Asian	N/A	80.61	N/A	82.99
American Indian	N/A	65.31	N/A	61.37

Note. The percent of student's proficient is calculated if the total number of students taking the test is 10 or greater. If fewer than 10 students took the test, the data will show as N/A in the data table, and will not be charted. Source: iowaschoolprofiles.com

Participants

This is a case study of 11 teachers of 75 second graders enrolled in a dual language program in an elementary school in a rural Midwestern town. The school had been identified for field testing of Spanish language directions for administration of the CogAT 7 because of its unique dual language program. The dual language program is a choice program, with about 60% of students in the district opting to participate. Native language literacy is emphasized and students learn together in mixed, linguistic groupings for literacy, math, science, and social studies in English and Spanish emphasizing a 50/50 model where approximately 50% of instruction is in English and 50% is in Spanish. The program strives for a balance between native English-speakers and native Spanish-speakers. At the elementary levels, instruction is also balanced. Students in the program receive half of academic instruction in English, and the other half in Spanish, with languages changing weekly. The dual language program in CSD has been active since 1998 and extends through the twelfth grade. Second grade was selected for this study as the talented and gifted program does not begin until third grade in CSD.

Participating teachers were drawn from a pool of 18 educators who had direct teaching experience with the student participants as primary dual-language classroom teachers (N=15) or as pull-out EL or Gifted and Talented teachers (N=3). Eleven teachers responded to a request for participation for a response rate of 61%.

The 75 students participated in the spring of their second grade year, when they took the Iowa Tests of Basic Skills (ITBS) Form A, primary achievement battery or the Logramos (Spanish-language equivalent) and the CogAT7 ability test battery with either Spanish or English directions, depending on home language. Teachers used professional judgment to determine the language most appropriate for students' testing.

Measures

CogAt 7

The CogAT Form 7 includes three batteries of tests: verbal, quantitative, and non-verbal (Lohman & Hagen, 2007). Though the batteries can be administered separately, all three were used for this study. The Primary Edition of the CogAT is typically administered for students in grades K-2 and requires no reading ability. The test administrator reads a question out loud and the children choose the picture that best answers the question.

The Iowa Tests of Basic Skills (ITBS)

This is an achievement test that is divided into batteries including vocabulary, word analysis, reading comprehension, listening, language, mathematics, science, sources of information, and social studies at levels 5-8, typically administered to students in grades K-2. The Core Battery contains all tests in the Complete Battery except Social Studies, Science, and Sources of Information. All tests at Levels 7 and 8, with the exception of Vocabulary and Reading, are orally administered.

Logramos, Second Edition

The Logramos, Second Edition is a group administered achievement test battery in Spanish. With content parallel to the ITBS, Logramos levels 7 and 8 are typically administered in first and second grades. Batteries include vocabulary, word analysis, reading, listening, language, math concepts, math problem solving, and math computation.

All tests at Levels 7/8, with the exception of vocabulary, reading, and a portion of mathematics computation, are orally administered. Students mark their answers in the machine-scorable booklets.

Procedures for Gathering Data

The research team worked closely with the district to facilitate administration of the CogAT Form 7 for second grade students. Test administrators read directions to students aloud and students choose the picture that best completes the problem. All three batteries (Verbal, Quantitative, Non-Verbal) were administered to students. Teachers in the bilingual program identified which students they felt should take the CogAT with directions in English and which students should take the CogAT in Spanish based on home language and teacher perception of ability. Of the students whose home language was Spanish (N=27), 82% were selected by teachers to receive CogAT directions in Spanish (N=22).

After CogAT administration and before receiving results of the ability test teachers were asked to rank order their students' academic abilities in reading and math based on classroom observations and teacher-built assessments. The research team mailed a questionnaire to teachers for completion. The teachers used a forced-choice format for ranking on the questionnaire, arranging equal numbers of students into groupings ranked high, middle, and low. The completed forms were submitted to an official at an intermediate education agency who connected teacher rankings of students with standardized test measures. This official removed any identifying student information before sending data to the researchers.

Researchers then ranked students based on performance on the standardized test measures, specifically the CogAT non-verbal core, CogAT quantitative core, CogAT verbal core, ITBS Reading Core, ITBS Math Core, Logramos Math Core, and Logramos Reading Core. National norms were used to code scores into high, medium, and low groups.

Quantitative Data Analysis Methods

Data were collected with the technical help of personnel from an intermediate education agency. The data were submitted directly to the intermediate education agency and were received from the agency with all identifying information removed. Data were transferred from Excel to SPSS. Results from the testing instruments were coded with a 1, 2, or 3 to indicate if the student's result was in the top, middle, or bottom tercile of scores. The teacher rankings of high, medium, and low ability were also coded with a 1, 2, or 3 and SPSS was used to run correlational studies. This study examines the relationships between the teacher perceptions of student ability and achievement and student ability and achievement as represented by standardized testing.

Results

Teacher Rankings

Teachers ranked students with a home language of English higher than those who did not speak English at home ($p = .01$). Students who were selected by teachers to receive directions in English on the CogAT were ranked higher than students who teachers selected to take the CogAT with Spanish directions for administration ($p=.05$). Table 3 outlines these results.

Table 3. Teacher Ranking, Home Language, and Language of Directions

	N	Average Teacher Ranking
Total	75	2.00
Home Language English	47	1.84*
Home Language Other	29	2.26*
English Directions on CogAT	55	1.90**
Spanish Directions on CogAT	20	2.27**

*Significant at $P=.01$; **Significant at $P=.05$

Teacher rankings, CogAT rankings, and ITBS/Logramos rankings were exactly the same for only 6% of students ($N=5$). Teacher rankings were positively correlated ($r = .269$, $p = .05$) with ability as measured by CogATVQN scores (Table 4). However, teacher rankings were more highly correlated with measures of achievement on the ITBS Reading ($r = .657$, $p = .01$), ITBS Math ($r = .602$, $p = .01$), Logramos Reading Comprehension ($r = .575$, $p = .01$), and Logramos Math ($r = .627$, $p = .01$).

Table 4. Teacher Ratings, Student Ability, and Student Achievement

		CogAT VQN	Teacher Rating	ITBS Reading	ITBS Math	Logramos Reading Comprehension	Logramos Math
CogAT VQN	Pearson Correlation	1	.269*	.156	.174	.381*	.432**
	Sig. (2-tailed)		.020	.166	.122	.011	.004
	N	81	74	80	80	44	43
Teacher Rating	Pearson Correlation	.269*	1	.657**	.602**	.575**	.627**
	Sig. (2-tailed)	.020		.000	.000	.000	.000
	N	74	74	74	74	40	39
ITBS Reading	Pearson Correlation	.156	.657**	1	.438**	.333*	.545**
	Sig. (2-tailed)	.166	.000		.000	.027	.000
	N	80	74	80	80	44	43
ITBS Math	Pearson Correlation	.174	.602**	.438**	1	.335*	.550**
	Sig. (2-tailed)	.122	.000	.000		.026	.000
	N	80	74	80	80	44	43
Logramos Reading Comprehension	Pearson Correlation	.381*	.575**	.333*	.335*	1	.494**
	Sig. (2-tailed)	.011	.000	.027	.026		.001
	N	44	40	44	44	44	43
Logramos Math	Pearson Correlation	.432**	.627**	.545**	.550**	.494**	1
	Sig. (2-tailed)	.004	.000	.000	.000	.001	
	N	43	39	43	43	43	43

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

Underestimation of Student Ability

By comparing teacher rankings with rankings on the CogAT researchers determined that teachers had either over estimated or under estimated abilities of 43% of students (N=28). Achievement was either overestimated or underestimated in 36% of students (N=27) as represented by both ITBS Reading and ITBS Math scores. As represented by Logramos scores, achievement was either overestimated or underestimated in 34% of students (N=13). Table 5 details the over and underestimation of ability and achievement by standardized measure.

Table 5. Overestimation and Underestimation of Ability and Achievement

	<u>Overestimation</u>		<u>Underestimation</u>	
	N	%	N	%
CogATVQN	15	23	13	20
ITBS Reading	9	12	18	24
ITBS Math	5	6	22	30
Logramos Reading	4	10	13	33
Logramos Math	1	3	12	31

Discussion

The research questions for this study were: to what extent do the teacher ratings of students and the performance on standardized tests of achievement and ability match?, what kinds of data are needed for teachers to make decisions about gifted and talented placement for children who are ELs?, and what talents are overlooked or undervalued using the data that are available to the teachers? We found that while teachers' perception of student ability and achievement were positively correlated with standardized test measures of ability and achievement, teachers were more accurate in perceptions of achievement than ability. Though some may think that enrollment in a dual language program may make assessments of teacher assessments of student ability more accurate, we found that multiple measures were necessary to gain a full understanding of a student's ability. Given that teachers were more likely to rank students whose primary language was English higher than those whose primary language was other than English, we call for the importance of using multiple measures when making decisions about placement in gifted and talented programs. Teacher rankings alone are insufficient to make placement decisions. This study highlights the need for proper assessment tools to identify gifted and talented language minority speakers.

Limitations and Pedagogical Implications

This study is limited by the potential bias inherent in teacher perceptions. Teachers chose which students were to take the ability tests with Spanish directions and which students were to take the test with English directions. Furthermore, the categorization of

students into high, medium, and low groups used broad classifications of ability and achievement. Future studies would gain more information by using a finer grain of classification.

Like many school districts around the country with a large, English-learning population, CSD finds itself with very few Hispanic high schoolers enrolled in advanced or AP math and science classes. By identifying the talents in the elementary children without the interference of language, there is hope that the long-term result will be that a greater percentage of Hispanic students will have the groundwork provided to them in elementary school be prepared for advanced coursework later in their academic careers. We also call for additional trainings for teachers to help understand and use results of standardized test measures, as well as information sessions for parents of ELs. Further research, including teacher interviews, would help complete a more in-depth understanding of teacher perceptions of ability and achievement.

Conclusions

Overall, the present findings provide evidence for the use of multiple data sources to make decisions about placement of students in gifted and talented programs. Teacher rankings alone are insufficient and unreliable means for evaluating the potential giftedness of EL students. Ford (1998) has argued that the misidentification of minority students for gifted and talented programs is a “tragic and unnecessary waste of human potential and promise” (p. 12). The results of this work can be used to inform a more socially just policy for the identification of and education of gifted and talented children learning English in a dual language program.

To get the full picture of what a student knows and is able to do, schools need multiple data sources. The rationale for this study is grounded in the growing critical need for appropriate, psychometrically sound tools to measure fairly what all school children know and are able to do. Specifically, there is a need for assessments that can provide educators with valuable information to allow for educational interventions for students to achieve their academic potential, particularly students from linguistic minority groups.

References

- Barkan, J. H., & Bernal, E. M. (1991). Gifted education for bilingual and limited English proficient students. *Gifted Child Quarterly*, 35(144).
- Bialystok, E. (2007). Cognitive Effects of Bilingualism: How Linguistic Experience Leads to Cognitive Change. *International Journal of Bilingual Education and Bilingualism*, 10(3), 210-223.
- Boyle, A., Taylor, J., Hurlburt, S., & Soga, K. (2010a). Title III accountability: Behind the numbers. *Evaluation brief*. Washington, DC: American Institutes for Research. Retrieved from http://www.air.org/files/Title_III_Behind_the_Numbers_043010.pdf.
- Boyle, A., Taylor, J., Hurlburt, S., & Soga, K. (2010b). Title III accountability: Behind the numbers. *Evaluation brief*. Washington, DC: American Institutes for Research.
- Callahan, C. M. (2005). Identifying gifted students from underrepresented populations. *Theory Into Practice*, 44(2), 98-104.
- Castellano, J. A., & Diaz, E. I. (2002). Reaching New Horizons: *Gifted and Talented Education for Culturally and Linguistically Diverse Students*, ERIC.

- Esquierdo, J. J., & Arreguin-Anderson, M. (2012). The "Invisible" Gifted and Talented Bilingual Students: A Current Report on Enrollment in GT Programs. *Journal for the Education of the Gifted*, 35(1), 35-47.
- Ford, D. Y. (1998). The underrepresentation of minority students in gifted education: problems and promises in recruitment and retention. *The Journal of Special Education*, 32(1), 4-14.
- Fultz, M., Lara-Alecio, R., Irby, B. J., & Tong, F. (2013). *The Hispanic Bilingual Gifted Screening Instrument: A Validation Study*. Texas A&M University.
- García, E., Jensen, B., & Scribner, K. (2009). The demographic imperative. *Educational Leadership*, 66(7), 8-13.
- Gonzalez, V. (2006). Profiles of Cognitive Developmental Performance in Gifted Children Effect of Bilingualism, Monolingualism, and Socioeconomic Status Factors. *Journal of Hispanic Higher Education*, 5(2), 142-170.
- Gonzalez, V. (2012). Assessment of Bilingual/Multilingual Pre-K–Grade 12 Students: A Critical Discussion of Past, Present, and Future Issues. *Theory Into Practice*, 51(4), 290-296.
- Hollingworth, L. (2008). Unintended educational and social consequences of the No Child Left Behind Act. *J. Gender Race & Just.*, 12, 311.
- Iowa Department of Education. (2007). *Educating Iowa's English Language Learners (ELLs): A Handbook for Administrators and Teachers*. Des Moines, IA.
- Lindholm-Leary, K., & Borsato, G. (2006). Academic Achievement. In F. Genesee, K. Lindholm-Leary, W. M. Saunders & D. Christian (Eds.), *Educating English language learners: A synthesis of research evidence* (pp. 176-211). New York: Cambridge University Press.
- Lohman, Gambrell, J., & Lakin, J. (2008). The commonality of extreme discrepancies in the ability profiles of academically gifted students. *Psychology Science Quarterly*, 50(2), 269-282.
- Lohman, D. F., Korb, K. A., & Lakin, J. (2008). Identifying Academically Gifted English-Language Learners Using Nonverbal Tests. *Gifted Child Quarterly*, 52(4), 275-296.
- Merriam, S. B. (1998). *Qualitative Research and Case Study Applications in Education*. San Francisco, CA: Jossey-Bass.
- Norman, D., & Balyeat, R. (1973). Whither ESEA III? *The Phi Delta Kappan*, 55(3), 190-192.
- Ramsey, A., & O'Day, J. (2010). Title III policy: State of the states. *ESEA evaluation brief: The English Language Acquisition, Language Enhancement and Academic Achievement Act (ESEA)*. Washington, DC: American Institutes for Research.
- Snow, R. E. (1977). Individual Differences and Instructional Theory. *Educational Researcher*, 6(10), 11-15.
- Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher*, 3-13.
- Spellings, M. (2005). *Speeches: Academic gains of English Language Learners prove high standards, accountability paying off*. Washington, D.C.: Retrieved from <http://www2.ed.gov/news/speeches/2005/12/12012005.html>.
- U.S. Department of Education. (2012). *National evaluation of Title III implementation: Report on state and local implementation*. Retrieved from www2.ed.gov/rschstat/eval/title-iii/state-local-implementation-report.pdf.
- Vald, G., & Valdes, G. (2003). *Expanding definitions of giftedness: The case of young interpreters from immigrant communities*. Routledge.
- Valdes, G., & Figueroa, R. A. (1994). *Bilingualism and Testing: A Special Case of Bias*. Norwood, NJ: Ablex Publishing Corporation.
- Yin, R. (2003). *Case Study Research Design and Methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.





Caracterización de directores de escuelas primarias del sur de Sonora

Characteristics of Principals at Elementary Schools in Sonora, Mexico

Ana Patricia Luna Gómez, Sonia Beatriz Echeverría Castro y José Javier Vales García

Instituto Tecnológico de Sonora

Información adicional sobre este manuscrito escribir a:

Ana Patricia Luna Gómez, anapatricia.lg@hotmail.com, Sonia Echeverría Castro, soniae@itson.edu.mx, José Javier Vales García, javier.vales@itson.edu.mx

Cómo citar este artículo:

Luna Gómez, A. P., Echeverría Castro, S. B., & Vales García, J. J. (2015). Caracterización de directores de escuelas primarias del sur de Sonora. *Educación y ciencia*, 4(43), 19-30.

Resumen

El propósito de esta investigación fue caracterizar a los directores de escuelas primarias del sur de Sonora e identificar posibles diferencias entre directores de escuelas de alto y bajo nivel académico. Es un estudio descriptivo, participaron 17 directores y se aplicó un cuestionario con preguntas abiertas, de escala y opción múltiple, distribuidos en siete apartados con un total de 178 reactivos, fundamentados en los criterios de escuelas de calidad. Los resultados muestran similitudes en las características de los directores, las diferencias identificadas como significativas son la cantidad de horas dedicadas a la jornada laboral, de reuniones para discutir temas como la convivencia escolar y el desempeño académico, que fueron superiores para los directores de escuelas con alto nivel académico.

Palabras clave: directores, escuelas de calidad, desempeño académico

Abstract

The purpose of this research was to characterize the headmasters of elementary schools in southern Sonora and identify possible differences between principals of high and low academic level. It is a descriptive study, involving 17 directors and there was a questionnaire applied with open questions of scale and multiple choices, distributed in seven sections with a total of 178 items according to quality standards from Mexican National Institute for Evaluation. Results show similarities in characteristics of the directors, the differences identified as significant are the number of hours devoted to the working hours of meetings to discuss issues such as school coexistence and student performance, which are higher for directors of schools with high academic level.

Keywords: headmasters, quality schools, academic performance

Introducción

Las personas que llegan a ejercer el puesto de director en una escuela de educación básica mexicana, no tienen una preparación oficial previa para ejercer el puesto de director y son elegidos por los puntos que acumulan en el sistema de escalafón, que se basa en parte en su antigüedad laboral, lo que lleva a que en la actualidad muchos directores sean profesores que se encuentran en la etapa final de su carrera, por lo tanto queda en el juicio de cada profesor el prepararse para su nuevo puesto como director (Badillo, 2005).

Es importante preguntarse, cuáles son las características de los directores en la región, sus años de experiencia en funciones directivas, en docencia, sus horas de dedicación a su jornada laboral como director, qué tanto se involucran en cuestiones pedagógicas y otros aspectos que identifican y diferencian a un director de otro, así como la asociación entre estas características y los resultados académicos de sus estudiantes.

Objetivos

1. Identificar las características de los directores de la región, según las variables: formación y experiencia, distribución de su tiempo laboral, trabajo colegiado, dirección de recursos, planeación estratégica y mejora.
2. Determinar si existen diferencias en cuanto a las características entre directores de escuelas primarias con estudiantes que obtuvieron alto y bajo puntaje en la prueba ENLACE de español y matemáticas de niños de sexto año.

Situación actual de la educación básica en México

El Estado mexicano tiene la obligación de garantizar condiciones suficientes para asegurar, sin discriminación alguna, el derecho de todas las personas a recibir una educación de calidad. Esto significa, al menos, que los estudiantes asistan de manera regular a la escuela, permanezcan en ella hasta concluir la escolaridad obligatoria y, en ese tránsito, logren aprendizajes relevantes para la vida. Existen, por supuesto, condiciones estructurales que representan retos para lograr garantizar el derecho a la educación, destacando la mayor vulnerabilidad de las poblaciones socioeconómicamente marginadas, indígenas y aisladas geográficamente (INEE, 2012).

Según el INEGI (2014), de las escuelas de educación básica (hasta secundaria) el 40.1% son primarias, de éstas según datos del INEE (2012) más de una quinta parte carece de agua y 20% de energía eléctrica; siete de cada cien no tiene baños y, el 2% no cuenta con aulas o salones de clases. Las primarias generales son las mejor dotadas (75% dispone de los cuatro servicios básicos), mientras que apenas la cuarta parte de las escuelas indígenas y 16% de las comunitarias cuentan con estos servicios.

El papel del director de escuela primaria

El director de escuela primaria tiene como función principal la de gestionar la escuela e involucrarse en los asuntos pedagógicos; es el responsable inmediato de administrar la prestación del servicio educativo conforme a las normas y lineamientos establecidos por la Secretaría de Educación Pública (SEP), como la primera autoridad

responsable del correcto funcionamiento, organización, operación y administración de la escuela y sus anexos, es quien tiene el control de que la aplicación del plan y los programas de estudio se efectúen conforme a las normas, lineamientos y las demás disposiciones e instrucciones que en materia de educación básica escolarizada establezca la SEP. También es el encargado de prever y organizar las actividades, recursos y apoyos necesarios para el desarrollo del plan y los programas de estudio. Así como dirigir y verificar dentro del ámbito de las escuelas que las actividades de control escolar, de extensión educativa y de servicios asistenciales se realicen conforme a las normas y lineamientos establecidos. Por último, le corresponde evaluar el desarrollo y los resultados de las actividades del personal a su cargo en la escuela, las aulas y la comunidad (Escamilla, 2006).

De esta forma, la función directiva es fundamental para la vida escolar, porque de ella depende la organización, planificación, implementación de las reformas, gestión de recursos, evaluación y el generar las condiciones para lograr la función educativa y por supuesto que los alumnos obtengan los resultados académicos esperados (Araiza, Magaña y Carrillo, 2014).

Capacitación formal para el puesto de director

Las competencias para el ejercicio de un puesto de director son diferentes a las requeridas para la docencia por lo que se requiere un proceso de formación y capacitación; al respecto, Escamilla (2006) sugiere un proceso que le permita iniciar en las funciones y una actualización permanente que incorpore los aspectos que se rescaten de la misma experiencia y práctica directiva.

Hasta hace poco los directores llegaban a su puesto mediante el puntaje obtenido en el escalafón, que tomaba en cuenta aspectos como la preparación académica del profesor, la asistencia a cursos de formación requeridos y los años de experiencia en la docencia, entre otros factores. En las categorías del escalafón no se incluía algún tipo de formación previa para la gestión escolar, en consecuencia los directores iniciaban su puesto sin ninguna capacitación orientándose con el apoyo de colegas, subdirector y su propio criterio ante el puesto (Aguilera, 2011). Actualmente, con la publicación de los lineamientos para la selección del personal con funciones de dirección las dinámicas para la asignación de los puestos está comenzando a modificarse, pero aún no está claro si se incluirá un proceso de capacitación y las características que éste tendría, o considerando que se evalúan los criterios para ese perfil y competencias de inicio ya se espera que las traigan consigo (DOF, 2015).

En el director se deposita la responsabilidad del cambio y transformación para la mejora de la calidad, éste se convierte en un protagonista fundamental que necesita tener la visión del cambio en la cultura escolar; la complejidad de las instituciones escolares, la tarea educativa y los múltiples requerimientos a los que tienen que dar respuesta las escuelas, requieren de un director con competencias para lograr un cambio eficaz, y una función directiva asociada a la necesidad de elevar el desempeño académico de sus estudiantes (Camacho, 2013).

El director tiene también un impacto positivo en el profesor, quien es el principal actor que está en contacto con los alumnos; Valenti (2009) al buscar los determinantes del desempeño académico de los estudiantes de primaria y secundaria de la República Mexicana, identificó que el profesor es un fuerte factor para el desempeño escolar de su grupo escolar, un mayor involucramiento de éste produce ganancias de hasta 10 puntos en las pruebas ENLACE para sus estudiantes, y a su vez el mayor compromiso del profesor es

potenciado cuando es acompañado de mayores capacidades por parte del director de la escuela.

En otro estudio se analizó el tipo de problema que tienen un grupo de directores, conforme el momento en la trayectoria de su puesto, muestran que en los primeros años en el puesto se presentan principalmente dificultades con el personal, insuficiencia de recursos financieros, la gestión de sí mismo, relación con el inspector y los alumnos y el Programa Escuela de Calidad. Los directores señalaron que los profesores no toman en serio su trabajo, llegan tarde, no se involucran con el trabajo de los niños en el aula, se sienten inconformes por lo cual existe una muy baja comunicación entre el director y el profesor. En el estudio también se muestra que existen directores que trabajan doble turno, en uno son directores y en el otro son docentes, mencionaron que el tener doble plaza crea un conflicto cuando el profesor que es director en otro turno quiere tomar el rol de director cuando no le corresponde (López, Slater y García, 2010).

García, Slater y López, (2010) realizaron una investigación con una muestra de diez directores, cinco mujeres y cinco hombres, se examinaron los problemas y situaciones que enfrentan los nuevos directores de educación primaria. Los resultados señalan que los problemas más comunes que tienen al principio de su función refieren a la carga de trabajo administrativo, la relación con los alumnos, padres de familia y profesores es un poco difícil de manejar en este primer momento de su trayectoria. Uno de los problemas que mencionaron haber tenido con los docentes es la dificultad para ejercer su autoridad.

En otros hallazgos encontrados en directores de escuelas de educación básica mexicana, Schmelkes (2005) señala que no se involucran en el aspecto pedagógico y no participan con los docentes en este ámbito, especialmente quienes tienen doble plaza o bien otras actividades administrativas asignadas, por lo tanto dedican poco tiempo a la administración. Agregado a que no reciben una capacitación formalmente establecida para ejercer su rol de director, muchos asumen una actitud de que lo que sucede en las aulas no es asunto suyo, por lo que no participan en los asuntos pedagógicos de su escuela. Se recomienda capacitar adecuadamente al director para que logre promover una mayor motivación, iniciativa y creatividad en todos los integrantes del colectivo escolar a través del ejercicio de liderazgo como director.

Metodología

Es una investigación descriptiva realizada en escuelas primarias en tres municipios del sur de Sonora.

Participantes

Se consideraron a 17 directores de escuelas primarias, seleccionados con base en los resultados de los estudiantes de sus escuelas en la prueba ENLACE 2011 y 2012. Se eligieron a las escuelas que estuvieran en sus puntuaciones en ambos extremos, los altos y bajos puntajes, ubicadas en zonas urbanas y rurales, y eran de tipo pública y privada.

Instrumento

Se elaboró un cuestionario con siete apartados, con base al Modelo de Gestión Educativa Estratégica del Programa Escuelas de Calidad, es de tipo semiestructurado con

preguntas abiertas, cerradas y tipo escala. A continuación se mencionan los apartados del cuestionario del director: a) datos generales con cinco reactivos; b) formación previa y actualización, con 14 ítems c) distribución del tiempo laboral, con 25 preguntas; d) trabajo colegiado, con 23; e) planeación estratégica, con 33 reactivos; f) dirección de recursos, con 27 ítems; y g) mejora, con 28 reactivos.

Procedimiento

Primero se realizó una revisión bibliográfica de estudios de la función del director, y los análisis de la situación actual de éstos en el contexto de las escuelas de calidad. Posteriormente, se elaboró una encuesta considerando las variables encontradas en la revisión bibliográfica. Enseguida, después de solicitar los permisos correspondientes, se realizó la aplicación de entrevistas en diferentes primarias del sur de Sonora. Se elaboró la base de datos, utilizando un paquete estadístico para las Ciencias Sociales, SPSS 15. Seguido de ello, se realizaron análisis de frecuencia de las variables en estudio para caracterizar a los directores, y buscar identificar similitudes y diferencias entre las características que presentan quienes están en la función directiva de escuelas con resultados bajos y altos en las pruebas ENLACE.

Resultados y discusión

A continuación se presenta la caracterización del director de escuela primaria del sur de Sonora. Se muestran tablas de frecuencia de las variables en estudio seguido de las comparaciones entre estas variables para las escuelas consideradas en los grupos de alto y bajo nivel académico a través de tablas de contingencias.

Al comparar la experiencia previa del director como docente entre las escuelas de nivel alto y bajo, son similares como grupo; en ambos tipos de escuelas se encuentran quienes tienen de 1 a 5 años, como los que han sobrepasado ya los 25, por lo menos hay un director en cada una de las categorías de años de experiencia docente (Ver Tabla 1).

Tabla 1. Frecuencia de directores según sus años de experiencia en la docencia

Años de experiencia	Nivel de escuela		Total
	Alta	Baja	
1 a 5 años	2	1	3
16 a 20 años	2	3	5
21 a 25 años	3	3	6
26 a 30 años	2	1	3
Total	9	8	17

Al contrastar la antigüedad de los directores y el desempeño de la escuela, se observa que cuando tienen entre 1 a 5 años, son similares las cantidades, pueden dirigir escuelas que obtengan alto o bajo nivel de desempeño académico, de entre los 11 a los 30 años aparecen en alto desempeño, pero ello puede deberse a la poca cantidad de directores con esta antigüedad en la muestra. Ya que el único director de más de 30 años, se encuentra en una escuela de bajo desempeño (Ver Tabla 2).

Tabla 2. Antigüedad en el puesto y desempeño de la escuela

Años	Nivel de escuela		Total
	Alta	Baja	
1 a 5	5	4	9
6 a 10	2	3	5
11 a 15	1	0	1
26 a 30	1	0	1
31 a 35	0	1	1
Total	9	8	17

La duración de la jornada semanal de los directores es variable, al contrastar estos datos (con prueba t) entre las escuelas de alto y bajo desempeño académico, se encontró una media de 35 horas para los del grupo alto y de 23 para los de bajo (Ver Tabla 3) con una diferencia significativa de .027, mostrando una desviación entre 8 y 10 horas.

Tabla 3. Media de la jornada laboral semanal por nivel desempeño

Categoría	N	Media	Desviación típ.	Error típ. de la media
Alta	9	35	8.614	2.871
Baja	8	23	10.714	3.788

Por ejemplo, se observa que un director de una escuela de nivel alto trabaja 48 horas, mientras que en nivel bajo no hay ninguno con esa cantidad de horas; a su vez, cuatro directores de escuelas de alto nivel trabajan 40 horas en comparación de uno en una escuela de nivel bajo. Un director de una escuela de nivel bajo y uno alto, trabajan 30 horas a la semana, tres directores de escuela de nivel alto y cuatro directores de escuela de nivel bajo trabajan 25 horas. Dos directores de escuela de nivel bajo trabajan ocho horas semanales, permisos y otras comisiones parecen estar relacionados con este tipo de dedicación. En general se encuentra que los directores de escuelas de alto desempeño trabajan más horas que los de bajo desempeño (Ver Tabla 4). Según distintos autores (Caminero, 2012; Miranda y Jiménez 2010), el éxito de una escuela proviene del compromiso y trabajo en equipo de todos y por supuesto del líder educativo, es decir el director, al tener un mayor números de horas dedicadas a su labor está claro que podrían obtenerse mejores resultados académicos.

Tabla 4. Horas de jornada laboral semanal del director y nivel de escuela

Horas	Nivel de escuela		Total
	Alta	Baja	
48	1	0	1
40	4	1	5
30	1	1	2
25	3	4	7
8	0	2	2
Total	9	8	17

Con referencia al nivel de escuela y los apoyos de inducción recibidos, en las escuelas que tienen un nivel académico alto la mayoría de sus directores recibieron de tres a cuatro recursos para su inducción al puesto, mientras que uno reporta que sólo tuvo uno. En las escuelas con un nivel de bajo desempeño académico hay un director que no recibió ningún tipo de inducción y la cantidad de elementos que recibieron se encuentran entre cero a cuatro, en general se recibió menos introducción al puesto en estas escuelas de bajo desempeño (Ver Tabla 5). Estos resultados pueden ser un indicador de que las escuelas de bajo nivel académico no cuentan por alguna razón con las mismas oportunidades para mejorar el ejercicio de la función del director.

Tabla 5. Apoyos recibidos al director para la inducción al puesto

Recursos	Nivel de escuelas		
	Alta	Baja	Total
0	0	1	1
1	1	1	2
2	0	2	2
3	4	2	6
4	4	2	6
Total	9	8	17

La percepción de la utilidad de la capacitación recibida por el director (previa o en el comienzo del ejercicio de su función) fue variable, algunos consideraron útil algunos recursos, en tanto que otros los percibieron como de poca o regular utilidad. El valor más alto que se podía obtener era de 15, pero el valor máximo observado fue de 10 (Ver Tabla 6), ningún director consideró que fueran de mucha ayuda, lo cual coincide con lo encontrado por Silva, Coral y Cordero (2009), que la poca capacitación que se la ha dado el director no le genera beneficio en las competencias pertinentes al ejercicio de sus funciones directivas.

Tabla 6. Rango de puntuaciones de utilidad percibida de los recursos de capacitación recibida por los directores

Rango de puntuaciones de Utilidad percibida	Frecuencia
0	1
1-2	3
3-5	7
6 - 10	6
Total	17

La utilidad de los recursos de capacitación recibidos entre escuelas de alto y bajo nivel académico (el puntaje máximo de utilidad pudo haber sido 15) coinciden con similares puntajes, dando como máximo una puntuación de nueve lo que se considera de utilidad regular en los recursos recibidos para ambos tipos de escuelas (Ver Tabla 7).

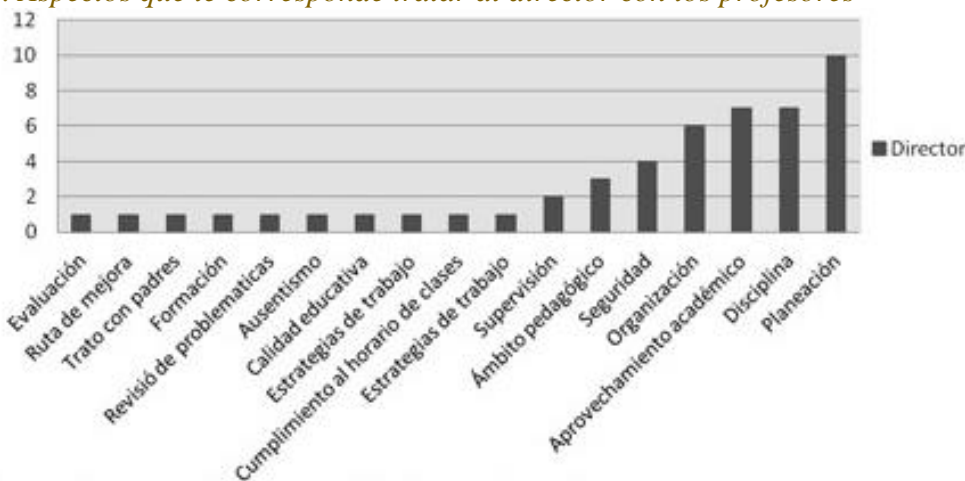
Tabla 7. Utilidad de recursos de capacitación recibidos y nivel académico de escuela

Utilidad	Nivel de escuela	
	Alto	Bajo
0		1
1-2	1	2
3-5	4	3
6-10	4	1
Total	9	8

Según los directores, los asuntos que más le competen trabajar con sus profesores son los de planeación, disciplina, aprovechamiento académico y organización; entre los menos señalados es la evaluación, mejora, trato con padres, revisión de problemáticas, ausentismo y las estrategias de trabajo.

Llama la atención que solo un director mencione la calidad como un tema a tratar con los docentes, si bien es cierto que este término puede englobarse en otros aspectos mencionados, es curioso que no se hable directamente del asunto de calidad educativa, cuando es un tema polémico, asociado directamente con la Reforma Educativa del 2013, y las manifestaciones de algunos grupos de docentes en protesta por ésta revelan que es un asunto de discusión entre este gremio de la educación (Ver figura 1).

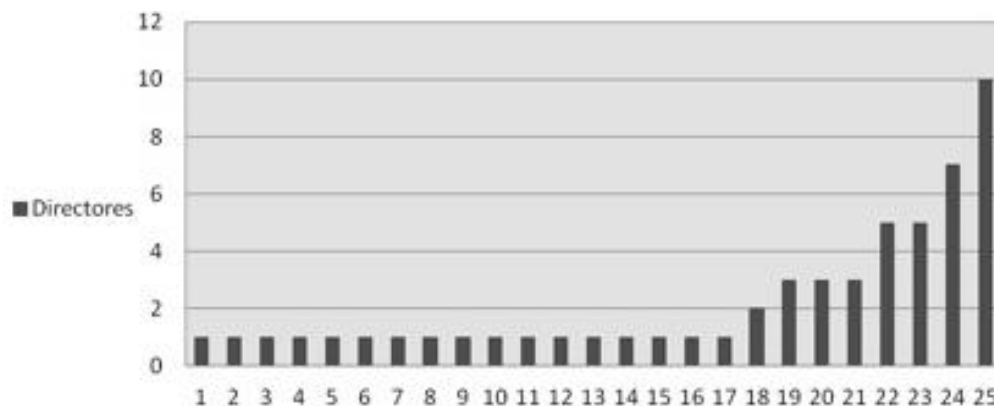
Figura 1. Aspectos que le corresponde tratar al director con los profesores



En la figura 2 se observa que los temas que más reportan los directores que se tratan en las reuniones con los padres de familia se refieren a la disciplina, el apoyo a los hijos y los referentes con el desempeño y las boletas. Entre los aspectos menos señalados se encuentran los relacionados con información acerca de proyectos, planeaciones del Plan Anual de Trabajo (PAT) y Plan de Apoyo Económico (PAE) y la mejora de la escuela; también el bullying o acoso escolar aparece con una frecuencia baja, lo cual es interesante cómo siendo un problema importante aparezca marginal ya que ni el director ni los padres de familia retoman el tema en las reuniones. Otros aspectos que se considerarían importantes a tratar en las reuniones con los padres por parte del director, son asuntos

financieros y las necesidades que presenta el plantel, así como la asistencia, sin embargo no aparecen como de los asuntos más tratados.

Figura 2. Temas más importantes tratados con los padres de familia



1	Información sobre proyectos	14	Bullying
2	Aclaración de dudas	15	Aprendizajes alcanzados
3	Planeación PAE Y PAT	16	Programación de actividades
4	Avance de metas planteadas	17	Sugerencias de apoyo académico conductual
5	Participación	18	Gestiones y acuerdos
6	Actividades escolares	19	Financieros
7	Organización escolar	20	Necesidades del plantel
8	Mejora en la escuela	21	Asistencia
9	Salud	22	Entrega de boletas
10	Reglamento escolar	23	Desempeño de alumnos
11	Educación	24	Apoyo a hijos
12	Salud	25	Disciplina
13	Seguridad		

Conclusiones

El tiempo invertido para ejercer sus funciones directivas parece marcar una diferencia significativa. Lo anterior es lógico considerando que para lograr realizar las acciones pertinentes es necesario tiempo; sin embargo se encontraron, variaciones grandes entre los directores, especialmente se distinguen que en la mayoría de las escuelas con bajo desempeño los tiempos de dedicación del director es menor que la cantidad de horas que trabaja el director en las instituciones de altos resultados. Una programación adecuada de los mismos y el seguimiento y trabajo con los problemas y mejoras de lo observado siempre requiere invertir tiempo.

En cuanto a las características de los directores, se concluye que tienen distintos años de experiencia como docentes y en cuanto al tiempo que llevan ejerciendo el puesto de director en la mayoría de los casos no tienen más de 10 años, es decir, hay profesores que tienen de poca a regular experiencia.

Los directores en su mayoría dijeron haber recibido algún elemento de apoyo como inducción al puesto de director, aun cuando en México no existe una preparación oficial

para ejercer el puesto de director, los directores señalaron que dichos elementos no fueron de gran utilidad para su preparación al puesto como director.

Los temas principales que los directores mencionaron tratar con los profesores son la planeación, disciplina, aprovechamiento académico y la organización. Entre los distintos temas que mencionaron, no aparece la discusión de la calidad como tal, lo que hoy es un tema importante para la educación en México.

Los temas que más directores coincidieron que tratan con los padres de familia fueron relacionados con la disciplina y el desempeño académico de los hijos, el apoyo que se les brinda a los alumnos y las asistencias de los mismos, también se trataron temas de carácter financiero del plantel escolar y otros aspectos de la escuela en general. Los padres de familia están interesados principalmente en discutir con el director sobre el desempeño escolar de sus hijos por lo que el director necesita tener una comunicación constante con los docentes para poder responder a las interrogantes de los padres de familia.

Recomendaciones

De los resultados se desprende la necesidad de tener una formación programada y evaluada para quienes aspiran al puesto de director, en las dos grandes áreas y sin afectación de las concepciones de funciones como las administrativas-gestión y las pedagógicas; ambas que se especifiquen sus alcances y lo que se espera de ese desempeño directivo, que los directores cuenten con una formación obligatoria y oficial por parte de la SEP previa al puesto del director, lo cual generaría directores más preparados para los retos administrativos y pedagógicos que representa el puesto de director de escuela básica.

Definir con exactitud y por escrito las responsabilidades que le corresponden a cada uno de los actores educativos en cuanto a su responsabilidad para el logro académico, y los asuntos que le corresponden tratar con los padres de familia, para tener una mejor comunicación y establecer con mayor claridad el compromiso por una corresponsabilidad en el logro del aprendizaje de los estudiantes.

Es importante señalar que la evaluación para ocupar puestos de director es actualmente una estrategia vigente, con la publicación de los lineamientos para la selección de personal con funciones a dirección (DOF, 2015), aún quedan retos importantes referidos al uso de esta información para la toma de decisiones, además de la calidad de los instrumentos aplicados, sin embargo, esa tendencia de incorporar la evaluación es necesaria para mejorar el Sistema Educativo Mexicano.

Referencias

- Aguilera, M. A. (2011). La función directiva en secundarias públicas. Matices de una tarea compleja. Instituto Nacional de Evaluación Educativa, Reporte de Investigación 35. Recuperado de: www.inee.edu.mx/images/stories/Lafunciondirectiva/lafunciondirectiva.pdf.
- Angulo, G. (2001). De Jomtien a Dakar. Diez años de ayuda a la Educación para Todos. Madrid: IntermónOxfam.
- Araiza, S., Magaña, R. & Carrillo, L. (2014) Evaluación por estándares de la gestión directiva en secundaria. Revista Iberoamericana de Educación. 64. Recuperado de: <http://www.rieoei.org/rie64a06.pdf>

- Badillo, M. (2005). El director como líder en la gestión educativa y el desafío para elevar la calidad de la educación básica, en M. Rueda Beltrán (Ed.), Congreso Nacional de Investigación Educativa. México: COMIE.
- Camacho, D. (2013). Competencias directivas para promover una gestión de calidad. Trabajo presentado en el primer Congreso Internacional de Transformación Educativa. Recuperado de <http://www.transformacion-educativa.com/congreso/ponencias/334-competencias-directivas.html>
- Caminero, J. (2012). Competencias de la dirección escolar para una gestión de calidad. Universidad de Valladolid. Recuperado de: <https://uvadoc.uva.es/bitstream/10324/1733/1/TFG-L5.pdf>
- DOF, (2015). Diario Oficial de la Federación. Recuperado de: http://www.dof.gob.mx/nota_detalle.php?codigo=5387781&fecha=03/04/2015.
- Escamilla, S. (2006). El director escolar. Necesidades de formación para un desempeño profesional. Tesis de doctorado. Universidad Autónoma de Barcelona.
- Figuerola, H. (2013). Reforma educativa no se negoció con CNTE para levantar plantón: Segob. Periódico excelsior especiales reforma educativa jaque a maestros versión en línea. Recuperado de: <http://www.excelsior.com.mx/nacional/2013/10/07/922220m>
- García, J., Slater, C. & López, G. (2010). El director escolar nivel primaria. Revista Mexicana de Investigación Educativa (15), (4), 1051.
- López, G., Slater, C. & García, J. (2010). Prácticas de Dirección y Liderazgo en las Escuelas Primarias Públicas de México. Los primeros años en el puesto. Revista Iberoamericana sobre calidad, eficacia y cambio en educación 8, 4.
- INEE (2012). La Educación en México: Estado actual y consideraciones sobre su evaluación. Recuperado de: <http://educacionyculturaaz.com/wp-content/uploads/2013/02/INEE-Educacio%CC%81n-en-Me%CC%81xico-Estado Actual.pdf>. INEGI (2014). Comunicado 076.- Censo de Escuela, Maestros y Alumnos de Educación Básica y Especial. Recuperado en: <http://www.comunicacion.sep.gob.mx/index.php/comunicados/marzo-2014/463-comunicado-076-censo-de-escuelas-maestros-y-alumnos-de-educacion-basica-y-especial>
- PEC (2010). Programa Escuelas de Calidad Modelo de Gestión Educativa Estratégica. Primera edición. Recuperado de: <http://basica.sep.gob.mx/pec/pdf/dprograma/MatGestModulo1.pdf>
- Miranda, F. & Jiménez, E. (2010). Gestión y calidad de la educación básica casos ejemplares de escuelas públicas mexicanas. SEB. Recuperado de <http://www.benv.edu.mx/pdf/librorgce.pdf>
- Secretaría de Educación Pública (2001). Programa Nacional de Educación 2001-2006. México: SEP. Recuperado de <http://planipolis.iiep.unesco.org/upload/Mexico/Mexico%20Programa%20nacional%20de%20educacion%2020012006.pdf>
- Schemelkes, S. (2005). La desigualdad en la calidad de la escuela primaria. Revista Latinoamericana de Estudios Educativos, (15). Recuperado de: http://benv.edu.mx/reforma_curricular/MATERIALES_CURSO_PANORAM_ACTUAL/la%20desigualdad%20en%20la%20calidad%20de%20la%20educacion%20primaria.pdf
- Silva, B., Coral, L. & Cordero, G. (2009). Las capacidades del director de educación primaria en México desde la perspectiva de los expertos. Ponencia en el X Congreso Nacional de investigación educativa, área 15 procesos de información.
- Valenti, G. (2009). Factores asociados al logro educativo. Un enfoque centrado en el estudiante, análisis y redacción. Facultad Latinoamericana de Ciencias Sociales.
- Velaz, C. (2005). Retos de la educación básica en América Latina. Madrid, España: EFCA, S.A





Influencia del estrés en el rendimiento académico de un grupo de estudiantes universitarios

Influence of Stress on the Academic Performance of a Group of College Students

Rodrigo Arturo Domínguez Castillo, Gladys Julieta Guerrero Walker y José Gabriel Domínguez Castillo

Universidad Autónoma de Yucatán

Información adicional sobre este manuscrito escribir a:

Rodrigo Arturo Domínguez Castillo, ra.dominguez@correo.uady.mx, Gladys Julieta Guerrero Walker, gwalker@correo.uady.mx, José Gabriel Domínguez Castillo, jg.dominguez@correo.uady.mx

Cómo citar este artículo:

Domínguez Castillo, R. A., Guerrero Walker, G. J., & Domínguez Castillo, J. G. (2015). Influencia del estrés en el rendimiento académico de un grupo de estudiantes universitarios. *Educación y ciencia*, 4(43), 31-40.

Resumen

La presente investigación tiene como objetivo describir la influencia del estrés en el rendimiento académico de estudiantes del sexto semestre de la Licenciatura en Educación que ofrece la Universidad Autónoma de Yucatán. Se aludieron diversas investigaciones relacionadas con el estrés académico, con el fin de conocer factores que generan estrés en el ámbito escolar. La muestra estuvo compuesta por 20 estudiantes de distintos sexos, que cursaban la Licenciatura en Educación. El alcance de la investigación fue de tipo descriptivo-correlacional, aplicándoles a los estudiantes el Inventario SISCO del Estrés Académico. Los resultados de este grupo de estudiantes muestran que el estrés no afecta de manera significativa el rendimiento académico, sin embargo, éste interfiere en algunas de las actividades académicas realizadas por los discentes.

Palabras clave: estrés académico, rendimiento académico, estudiantes universitarios, factores estresores

Abstract

The main objective in this article is to describe the influence of stress on the academic performance. The article points out various research projects related to the academic stress in order to understand factors that generate stress at school level. The sample was composed of 20 students from different sexes that were enrolled in the sixth semester in the Education Program at the Universidad Autonoma de Yucatan. The scope of the research was descriptive and correlational study. Students responded the SISCO inventory of academic stress. The results of this group of students show that the stress does not significantly affect the academic performance; however, it interferes with some of the academic activities conducted by the students.

Keywords: academic stress, academic achievement, university students, stress factors

Introducción

El concepto de estrés se remonta desde el siglo XVII con los descubrimientos de Robert Hooke y Thomas Young en el área de la física, mismos que permiten dar un nuevo enfoque al término a partir del análisis de éste desde la perspectiva de otras ciencias (Álvarez, 1989 citado en Román y Hernández, 2011).

A partir de esto, se desarrollan una variedad de definiciones que explican qué es el estrés, no obstante, quien acuñó el término fue Hans Seyle en 1936 (García y Orozco, 2004 citado en Tolentino, 2009), denominándolo como “un conjunto coordinado de reacciones fisiológicas ante cualquier forma de estímulo nocivo” (p.4).

El estrés ha sido una temática de interés para diversos autores que intentan investigar los efectos que produce éste en los aspectos psicológicos, sociales y comportamentales de la vida del ser humano. Tal es el caso de Evans y Coben (1987) citado en Valdez (1997) que realizan estudios para determinar el impacto del estrés en las personas, clasificando éste en cinco áreas: efectos fisiológicos, efectos en la ejecución de tareas, efectos en las conductas interpersonal y afectiva, efectos en la conducta verbal y no verbal, y efectos en los procesos de adaptación.

El problema del estrés afecta a las personas sin hacer distinción alguna en el sexo, situación económica, edad o lugar de procedencia, es más Caldera, Pulido y Martínez (2007) afirman que éste se encuentra presente en todos los medios y ambientes, tales como en la familia, trabajo, escuela, relaciones sociales, etcétera.

Asimismo, existen una diversidad de factores llamados estresores que propician el desenvolvimiento del estrés, los cuales se pueden clasificar de la siguiente manera: factores psicosociales, bioecológicos y de personalidad (Lopategui, 2000).

Situación problemática

En cuanto al estrés, se han encontrado estadísticas que hacen referencia al nivel de afectación de éste en la sociedad. Se calcula que el 25% de la población mundial padece de estrés o de algún trastorno de la salud relacionado con éste, por lo cual se le considera uno de los principales problemas de salud psicológica en el mundo (Valdez, 1997 citado en Farkas, 2002).

Por otra parte, los escasos trabajos sobre el tema han demostrado la existencia de índices notables de estrés en las poblaciones universitarias, alcanzando mayores cuotas en los primeros cursos de carrera y en los períodos inmediatamente anteriores a los exámenes (Muñoz, 1999 citado en Martín, 2007) y siendo menor en los últimos. Lo anterior se relaciona al estrés académico, donde diversas situaciones afectan al estudiante.

La etapa de estudios universitarios representa el punto culminante del estrés académico por las altas cargas de trabajo, debido a que coinciden con una fase de cambios en la vida del estudiante (Arnett, 2000 citado en Pulido, Serrano, Valdés, Chávez, Hidalgo y Vera, 2011), también existen diversas situaciones que resultan estresantes para el estudiante como lo son los exámenes, los proyectos y trabajos finales los cuales repercuten en su desempeño académico.

Planteamiento del problema

De acuerdo con lo anterior, es necesario el estudio del estrés para conocer más a fondo dicho fenómeno. Según Román y Hernández (2011) el ambiente escolar es uno de los más estudiados, puesto que resulta un medio desencadenante de estrés, tomando en cuenta los diversos niveles de enseñanza.

De igual manera, se ha encontrado que los estudiantes son más propensos a presentar diversas situaciones estresantes que pueden afectar su desempeño, por lo tanto, en esta investigación se pretende estudiar la relación del estrés con el rendimiento académico de estudiantes universitarios; específicamente en el sexto semestre de la Licenciatura en Educación de la Facultad de Educación de la Universidad Autónoma de Yucatán.

Objetivos

Describir las situaciones que generan estrés en el ámbito escolar.
Determinar la influencia del estrés en el rendimiento académico.

Propósitos

Se pretende que la investigación profundice en la temática del estrés académico, determinando las situaciones estresantes así como la influencia de las mismas en el rendimiento académico de los estudiantes del sexto semestre de la Licenciatura en Educación de la Universidad Autónoma de Yucatán.

Definición de Variables

Estrés:

Definición constitutiva: es aquél que se produce a partir de las demandas que exige el ámbito educativo (Caldera, Pulido y Martínez, 2007).

Definición operacional: frecuencia de actividades académicas que generan en los estudiantes, sentimientos de inquietud y ansiedad.

Rendimiento académico:

Definición constitutiva: nivel de conocimientos, habilidades y destrezas que el alumno adquiere durante el proceso enseñanza-aprendizaje (Caldera, Pulido y Martínez, 2007).

Definición operacional: la calificación del semestre inmediato anterior, tomando en cuenta que el mínimo aprobatorio es 70, lo cual se consideró como un desempeño regular. El promedio 80, un desempeño bueno y 90 un desempeño excelente.

Justificación

Actualmente, la sociedad está viviendo una diversidad de transformaciones en su estructura social, política, económica y cultural debido a las nuevas exigencias del siglo XXI a las que se está enfrentando. Por la misma razón, los comportamientos de las personas se están modificando para poder adaptarse a las situaciones o demandas de su medio, generando significativamente condiciones estresantes.

El estrés es un término de carácter social que inconscientemente las personas utilizan a diario para hacer referencia a las reacciones físicas o psicológicas de tensión, ansiedad y cansancio, aunque autores como Caldera, Pulido y Martínez (2007) mencionan que se puede caracterizar como “todo tipo de amenaza que afecte directamente a un individuo” (p. 78).

Este fenómeno afecta a nivel mundial, puesto que uno de cada cuatro individuos sufre de este padecimiento o de algún problema de salud relacionado con el mismo (Valadez, 1997 citado en Farkas, 2002).

En el caso particular de México, la Universidad Sussez realizó un informe en el 2001 en el cual señala que éste es uno de los principales países con mayores niveles de estrés dado a la constante pobreza que vive (Caldera y Cols. 2007).

El estrés afecta en el ámbito personal, social, familiar y escolar de las personas, sin embargo, en este último, algunas de las repercusiones que podrían afectarlo son los aspectos fisiológicos, en la ejecución de tareas, en la conducta verbal y no verbal y en los procesos de adaptación (Evans y Cohen, 1987 citados en Farkas, 2002).

Por esta razón, en este estudio se hará énfasis en el estrés situado en el contexto académico dado a que, tal y como mencionan López y Díaz (s.f) “la población estudiantil es un grupo vulnerable propenso a periodos particularmente estresantes” (p.2); ya que a medida que el estudiante progresa en sus estudios, el estrés aumenta y alcanza los niveles más altos cuando éstos se encuentra en el nivel superior (Putwain, 2007; Dyson y Renk, 2006 citado en Pulido y Cols. 2011).

De igual forma, los diversos cambios que enfrentan al adaptarse a la vida universitaria y las decisiones que deben de tomar en su vida personal respecto a su proyecto de vida y carrera son situaciones que generan estrés a los estudiantes.

Es necesario que las instituciones detecten aquellos estudiantes afectados por el estrés para poder realizar acciones reactivas con el fin de apoyar a los alumnos en el transcurso de sus estudios, con el objetivo principal de evitar el fracaso escolar y el bajo rendimiento académico.

Revisión de la Literatura

El estrés ha sido uno de los temas estudiados actualmente, ya que se presenta en niños, jóvenes y adultos, influye en la vida del ser humano y afecta el comportamiento y desenvolvimiento de éste en la sociedad. No obstante, diversos autores han centrado sus investigaciones para profundizar en dicha temática, haciendo énfasis en las repercusiones de éste en el rendimiento académico.

Por ello, es indispensable poner atención en este problema en la educación, pues es en este ámbito donde se encuentran casos de ansiedad por trabajos escolares, exámenes parciales y finales, existe la presión en trabajos de equipo, el acoplamiento a diversos ambientes de trabajo, entre otros.

Es en la universidad, en la que la mayoría de los jóvenes padecen de esta problemática, pues estos mismos se preparan para la profesión que en un futuro ejercerán, por ende, el nivel de exigencia es mayor y aumenta la responsabilidad. Sin embargo, algunos autores mencionan que existe mayor nivel de estrés en el primer año de la carrera debido a la falta de control, no obstante, este será superado en el transcurso del tiempo en la universidad” (Tolentino, 2009; p. III).

El estrés es una respuesta inesperada del organismo ante situaciones agobiantes (Sandín, 1995; Keith, 1994 citado en Barraza, 2011). Por ello, el estrés aumenta cuando el nivel de exigencia en semestres avanzados es creciente. Aunado a esto, se encuentran las épocas de exámenes y trabajos en clase que generan ansiedad en los alumnos.

Como se mencionó anteriormente, el estrés no solamente afecta a jóvenes y adultos, del mismo modo, se presenta en estudiantes de primaria, secundaria y preparatoria (Connor, 2003, Aherne, 2001, Gallagher y Millar, 1996, citado en Pulido y Cols. 2011). Esto demuestra que el estrés no hace discriminación alguna entre edades, sexo, nivel socioeconómico, etcétera.

Arnett (2000) citado en Pulido, y Cols. (2011) coincide con lo que menciona Tolentino (2009), señalando que el estrés académico alcanza sus estándares más altos en los estudios de nivel superior, ya que estos se caracterizan por altas cargas de trabajo, los cuales son solicitados con altos niveles de exigencia y calidad.

Por esta razón, varios estudiantes han llegado a considerar la posibilidad de darse de baja de la institución debido al estrés académico que originan las altas cargas de tareas, actividades, evaluaciones, proyectos, etcétera (Pulido, y Cols. 2011).

Relacionado el estrés con el rendimiento académico, Caldera, y Cols. (2007) afirman que los altos niveles de estrés pueden considerarse como obstáculos para el rendimiento académico, por lo que se recomienda mantener un nivel medio de éste para incrementar la posibilidad de obtener buenas notas. No obstante, es necesario conocer otras causas de estrés en los estudiantes, pues existen otros motivos que generan ansiedad en los mismos y no solamente aspectos relacionados con la educación.

Uno de los principales generadores de estrés en los estudiantes es la evaluación académica (Farkas, 2002). Otro tipo de estresor, puede ser causado por los mismos maestros, debido a que estos exigen más a ciertos estudiantes y no al grupo en su totalidad (Barraza, 2011).

El estudio de Zeidner (1992) citado Kadapatti y Vijayalaxmi (2012) informó que la sobrecarga de tareas y los procedimientos de evaluación académica genera más presión y estrés en los estudiantes que los factores familiares, sociales y personales de los mismos.

Es necesario poner atención a las consecuencias del estrés en los estudiantes, puesto que como se mencionó anteriormente éste afecta a niños, adolescentes, jóvenes y adultos, deteriorando el estado emocional, la salud física o las relaciones interpersonales de los mismos (Martin, 2007).

Metodología

La metodología en la que se sustenta esta investigación es descriptiva y correlacional, ya que la intención se centra en describir las situaciones que generan estrés en el ámbito escolar y la influencia de éste en el rendimiento académico. Las variables a medir son: estrés escolar y rendimiento académico.

La muestra estuvo conformada por 20 estudiantes del sexto semestre de la Licenciatura en Educación de la Facultad de Educación impartida en la Universidad Autónoma de Yucatán. El tipo de muestreo que se utilizó fue no probabilístico.

Para reunir los datos, se administró el inventario SISCO del Estrés Académico, el cual tiene como objetivo recabar información sobre las características del estrés que suele acompañar a los estudiantes de educación media superior, superior y de postgrado durante

sus estudios. Dicho cuestionario fue modificado para fines de la investigación, anexando algunas preguntas relacionadas con el rendimiento académico, así como algunas de las actividades escolares afectadas por el estrés.

El Inventario SISCO del Estrés Académico se configura por 31 ítems distribuidos de la siguiente manera:

-Un ítem de filtro que, en términos dicotómicos (sí-no) permite determinar si el encuestado es candidato o no a contestar el inventario.

-Un ítem que, en un escalamiento tipo Likert de cinco valores numéricos (del 1 al 5 donde uno es poco y cinco mucho) permite identificar el nivel de intensidad del estrés académico.

-Ocho ítems que, en un escalamiento tipo Likert de cinco valores categoriales (nunca, rara vez, algunas veces, casi siempre y siempre) permiten identificar la frecuencia en que las demandas del entorno son valoradas como estímulos estresores.

-15 ítems que, en un escalamiento tipo Likert de cinco valores categoriales (nunca, rara vez, algunas veces, casi siempre y siempre) permiten identificar la frecuencia con que se presentan los síntomas o reacciones al estímulo estresor.

-Seis ítems que, en un escalamiento tipo Likert de cinco valores categoriales (nunca, rara vez, algunas veces, casi siempre y siempre) permiten identificar la frecuencia de uso de las estrategias de afrontamientos.

-Tres preguntas relacionadas con el rendimiento académico. La primera pregunta se refiere al promedio inmediato anterior del estudiante; la segunda, de tipo dicotómica (sí-no) hace referencia a la existencia de afectación del estrés en el rendimiento académico; y la última, presenta seis opciones de actividades escolares en donde el estrés afecta el rendimiento académico (modificación realizada).

Resultados

El 100% de los sujetos que contestaron el instrumento externaron tener momentos de preocupación y nerviosismo durante el transcurso del semestre que se encuentran cursando. En cuanto al nivel de preocupación o nerviosismo de los estudiantes, el promedio fue de 3.20 tomando en cuenta que el instrumento maneja una escala del 1 al 5.

Las principales situaciones que generan inquietudes en los estudiantes son: el 70% de los encuestados señalaron que la sobrecarga de tareas y trabajos escolares casi siempre les generan inquietudes, por otra parte, el 75% menciona que algunas veces la personalidad y el carácter del profesor.

El 55% afirman que las evaluaciones de los profesores (exámenes, ensayos, trabajos de investigación, etc.) casi siempre les generan inquietud.

El 50% de los encuestados señaló que la participación en clase algunas veces les genera inquietud. Por otra parte, contar con tiempo limitado para realizar trabajos genera casi siempre inquietudes con un 50%.

En cuanto a las reacciones físicas que tuvieron los estudiantes en momentos de preocupación o nerviosismo, los dolores de cabeza o migrañas les afectan algunas veces al 50% de los encuestados. Las demás reacciones físicas no fueron tan significativas ni frecuentes.

Los problemas de concentración son la principal reacción psicológica que perturban a los estudiantes, al ser mencionada por un 55% de éstos y en una escala de algunas veces. Las demás reacciones psicológicas no fueron tan significativas ni frecuentes.

El 55% de los estudiantes encuestados señalaron que algunas veces aislarse de los demás es una reacción comportamental que enfrentan en momentos de preocupación o nerviosismo, asimismo, el 50% afirmó que algunas veces presentan desgano para realizar labores escolares.

Respecto a las estrategias que utilizan los estudiantes para enfrentar las situaciones de estrés, el 50% mencionó que la asertividad y la elaboración de un plan en la ejecución de tareas las utilizan como medio de afrontamiento.

Como puede observarse en la tabla 1, el 55% de los estudiantes tienen un promedio mayor a 90 en la calificación del semestre inmediato anterior, el 40% de los mismos cuentan con un promedio entre 80 y 89 y el 5% restante de los encuestados ubican su promedio por debajo de 79. Es necesario destacar que el promedio mínimo aprobatorio es de 70.

Tabla 1. Rendimiento académico del semestre inmediato anterior de los estudiantes encuestados

Rendimiento académico	Porcentaje de los encuestados
Alto (Mayor a 90)	55%
Regular (Entre 80 y 89)	40%
Bajo (Menor de 79)	5%

En cuanto a las actividades que los estudiantes consideran que el estrés afecta en su rendimiento académico se encontraron los siguientes resultados:

Tabla 2. Actividades que los estudiantes consideran que el estrés afecta su rendimiento académico

Actividades	Porcentajes
Tareas	35%
Proyectos finales	50%
Concentración durante clases	50%
Exámenes	20%
Calificaciones finales	10%
Otras actividades	10%

Como puede observarse, las actividades relacionadas con la elaboración de proyectos finales y la concentración durante las clases son las que generan estrés en los estudiantes. Por otra parte, el 70% de los encuestados afirmaron que el estrés sí repercute en el rendimiento académico.

Conclusiones

Se puede concluir que a pesar de que el 70% de los encuestados expresaron que el estrés afecta su rendimiento académico, éste no se ve reflejado en las calificaciones de los mismos, debido a que el 55% afirmaron poseer un rendimiento alto, un 40% un rendimiento regular y sólo el 5%, un rendimiento académico bajo.

En cuanto a las situaciones que generan inquietudes en los estudiantes, se encontró la carga de tareas y las evaluaciones; no obstante, las actividades en donde el estrés alcanza sus mayores niveles de influencia son los proyectos finales y la concentración durante las clases.

A pesar de que no hay una influencia directa del estrés en el rendimiento académico, es necesario que los profesores tomen en cuenta que la carga de trabajos finales generan en los alumnos altos niveles de estrés, por lo tanto, éstos deben planificar equitativamente las actividades a realizar a lo largo del semestre

Asimismo, el Centro de Orientación puede desarrollar talleres o actividades en el que los estudiantes puedan manejar la ansiedad, las inquietudes y preocupaciones generadas por el estrés, con el objetivo de evitar que los alumnos consideren darse de baja de la institución por no saber manejar o controlar adecuadamente su estrés.

Referencias

- Barraza, A. (2011). *Estrés, Burnout y bienestar subjetivo*. (1a edición). México: Red Durango de Investigadores Educativos, A.C.
- Caldera, J., Pulido, B., y Martínez, M. (2007). Niveles de estrés y rendimiento académico en estudiantes de la carrera de Psicología del Centro Universitario de Los Altos. *Revista Educación y Desarrollo*, 7, pp. 77-82. Recuperado de http://www.cucs.udg.mx/revistas/edu_desarrollo/antiores/7/007_Caldera.pdf
- Farkas, C. (2002). Estrés y afrontamiento en estudiantes universitarios. *Revista Psykhe*, 11 (1), pp. 57-68. Recuperado de <http://www.psykhe.cl/index.php/psykhe/article/viewFile/456/435>
- Kadapatti, M. y Vijayalaxmi, A. (2012). Stressors of academic stress. A study on pre-university. *Indian J.Sci.Res*, 3 (1), pp. 171-175.
- Lopategui, E. (2000). *Estrés: concepto, causas y control*. Recuperado de <http://www.saludmed.com/Documentos/Estres.html>
- López, A. y Díaz, I. (s.f.). Análisis estadísticos de las causas de estrés en estudiantes universitarios. Diagnóstico para acciones tutoriales. *Ponencia presentada en XI Congreso Nacional de Investigación Educativa*. Recuperado de http://www.comie.org.mx/congreso/memoriaelectronica/v11/docs/area_16/1843.pdf
- Martin, I. (2007). Estrés académico en estudiantes universitarios. *Apuntes de Psicología*, 25 (1), pp.87-99. Recuperado de http://www.cop.es/delegaci/andocci/files/contenidos/VOL25_1_7.pdf
- Pulido, M., Serrano, M., Valdés, E., Chávez, M., Hidalgo, P. y Vera, F. (2011). Estrés académico en estudiantes universitarios. *Revista Psicología y Salud*, 21 (1), pp. 31-37. Recuperado de <http://www.uv.mx/psicysalud/psicysalud-21-1/21-1/Marco-Antonio-Pulido-Rull.pdf>
- Román, C. y Hernández, Y. (2011) El estrés académico: una revisión crítica del concepto desde las ciencias de la educación. *Revista Electrónica de Psicología Iztacala*, 14 (2), pp. 1-14. Recuperado de <http://www.revistas.unam.mx/index.php/rep/rep/article/view/26023>

- Tolentino, S. (2009). *Perfil de estrés académico en alumnos de licenciatura en Psicología, de la Universidad Autónoma de Hidalgo en la Escuela Superior de Actopan*. Tesis para obtener el título de Licenciada en Psicología, Escuela Superior de Actopan Universidad Autónoma del Estado de Hidalgo. Recuperado de http://www.uaeh.edu.mx/nuestro_alumnado/esc_sup/actopan/licenciatura/Perfil%20de%20estres%20academico%20en%20alumnos.pdf
- Valadez, A. (1997) Efectos del estrés: un análisis descriptivo. *Revista Interamericana de Psicología*, 32 (1), pp. 73-94. Recuperado de <http://dialnet.unirioja.es/servlet/articulo?codigo=3852501>





Relación entre aprendizaje de conceptos y aprendizaje de procedimientos en economía de la empresa agraria

A Relationship between Learning of Concepts and Learning of Procedures in the Agricultural Economics

Luis Alberto Araujo Andrade y Eraclio del Jesús Cruz Pacheco
Universidad Autónoma de Yucatán

Información adicional sobre este manuscrito escribir a:

Luis Alberto Araujo Andrade, aandrade@uady.mx, Eraclio del Jesús Cruz Pacheco, cpacheco@uady.mx

Cómo citar este artículo:

Araujo Andrade, L. A., & Cruz Pacheco, E. J. (2015). Relación entre aprendizajes de conceptos y aprendizaje de procedimientos en economía de la empresa agraria. *Educación y Ciencia*, 4(43), 41-53.

Resumen

Se estableció con base en la taxonomía revisada de Bloom, la incidencia del aprendizaje de conocimientos simples (conceptos) sobre el aprendizaje de conocimientos complejos (procedimientos) entre alumnos del programa de Licenciado en Economía de la Universidad Autónoma de Yucatán, organizados en grupos de trabajo permanentes durante el proceso enseñanza-aprendizaje de la Unidad V Economía de la Empresa Agraria de la asignatura Economía Agropecuaria, del cuarto semestre de dicho programa académico durante el período enero-junio de 2011. Los resultados revelan que se encontró una relación directa en el sentido de que a mayor grado de aprendizaje de conceptos (conocimiento simple) se presentó un mayor grado de aprendizaje de procedimientos en los procesos cognitivos de recordar, comprender y aplicar.

Palabras clave: aprendizaje, economía, empresa, conceptos, procedimientos

Abstract

This investigation is reviewing the incidence of learning of simple knowledge (concepts) on learning of complex knowledge (procedures) according to the Taxonomy of Bloom among students of Economics program at Universidad Autonoma de Yucatan. Students from the fourth semester were organized on permanent working groups during teaching-learning process of Unit V in the Agricultural Economics subject during the period January-June 2011. The output shows that there is a direct relationship in the sense that while students acquire a better level in the learning of concepts, they get a better understanding in learning of procedures on the cognitive process of remembering, comprehension and application.

Keywords: learning, economics, farm, concepts, procedures

Introducción

La educación superior, al basarse en el método científico, demanda el conocimiento de conceptos y procedimientos como parte de estructuras teóricas que permitan explicar una parcela de la realidad. En consecuencia, la enseñanza y el aprendizaje de los conceptos y los procedimientos son componentes fundamentales en la formación de los profesionistas.

De hecho el aprendizaje de los conceptos es el que le da sentido a las teorías, pues tales componentes, dentro de la jerarquía teórica a la que pertenecen, son los que presentan las conexiones lógicas entre las variables y que permiten interpretar los fenómenos de la realidad a la luz de una perspectiva científica y, por lo tanto, permiten la adquisición de aprendizaje significativo.

El marco teórico sobre el cual descansa el contenido de este artículo es, por un lado, la estructura jerárquica de la taxonomía de los objetivos de la educación de Benjamín Bloom y de sus revisores; y por otro lado el aprendizaje significativo y el aprendizaje de conceptos y procedimientos.

La investigación se llevó a cabo durante el período lectivo de enero a junio de 2011 y se refiere a la impartición de los contenidos de la Unidad V Economía de la Empresa Agraria que forma parte de la asignatura Economía Agropecuaria que corresponde al cuarto semestre de la Licenciatura en Economía de la Facultad de Economía de la Universidad Autónoma de Yucatán, en Mérida.

En consonancia con lo anterior, el objetivo del presente artículo es determinar de qué manera el aprendizaje de conceptos incidió en el aprendizaje de procedimientos en un campo del conocimiento científico, la economía de la empresa agraria. Éste es un estudio cuyos supuestos y explicaciones se enmarcan en las teorías constructivista y cognoscitiva de la educación, las cuales por medio de sus premisas proporcionan los elementos para interpretar los resultados a los que se llegó en esta investigación.

Taxonomía de los objetivos de aprendizaje y su revisión

Según Benjamín S. Bloom (1986), el dominio cognoscitivo o conocimiento es el que se refiere a “los comportamientos de recordar, razonar, resolver problemas, formar conceptos y, aunque de manera limitada, pensar creativamente” (p.16).

Según este autor “las conductas más complejas incluyen otras más simples” (p. 16). Conductas simples son aquellas que implican un esfuerzo mental reducido para poder aprenderlas, en comparación a las conductas complejas que implican un esfuerzo mental mayor (que las simples) para lograr su aprendizaje y las cuales no podrían ser aprendidas si no se posee ya un aprendizaje previo, que podría ser precisamente una o más conductas simples.

Las categorías de dominio cognoscitivo según la taxonomía de Bloom son las siguientes, de acuerdo al orden jerárquico de aprendizaje simple al más complejo: a) Conocimiento; b) Comprensión; c) Aplicación; d) Análisis; e) Síntesis; y f) Evaluación.

En el caso de las revisiones a la Taxonomía de Bloom, éstas se refieren a la incorporación de recientes aportaciones de la psicología del aprendizaje para superar el riguroso sistema jerárquico de dicha taxonomía debido a que se han encontrado contradicciones con el traslape entre varios objetivos educativos. Así, por ejemplo, la

comprensión de algunos conocimientos pueden ser de mayor complejidad que algunas actividades de análisis o de evaluación (Amer, 2006).

Aprendizaje significativo

La teoría del aprendizaje significativo indica que habrá aprendizaje cuando lo que se trata de aprender se relaciona de forma sustantiva con lo que ya conoce quien aprende, con aspectos relevantes y preexistentes de su estructura cognitiva. Un aprendizaje es significativo cuando el aprendiz puede atribuir posibilidad de uso (utilidad) al nuevo contenido aprendido relacionándolo con el conocimiento previo (Rivera, 2004).

Asimismo se considera al aprendizaje como construcción de significados, orientado a metas, situado, colaborativo, acumulativo, autorregulado e individualmente diferente. Bajo el enfoque constructivista, el sujeto interpreta la experiencia que elabora y prueba con base en conocimientos previos y conceptos compartidos; de esa manera el significado depende de la situación y convenciones sociales comunicadas (Rodríguez, Padilla y Esquivel, 2007).

Para aprender significativamente, los alumnos tienen que hacer algo con el conocimiento, construir relaciones entre la nueva información y sus conocimientos previos, utilizar el nuevo conocimiento para interpretar nuevas situaciones, hacer inferencias sobre relaciones potenciales, formular predicciones, resolver problemas, descubrir relaciones subyacentes, razonar y aprender a aprender (Cruz y Anzaldo, 2010).

Conceptos

Los conceptos son formas lógicas que caracterizan a toda una clase de objetos agrupados por uno u otro rango o como reglas de clasificación que permiten agrupar los objetos según ciertas características (Castañeda, 1998).

Los conceptos llevan a cabo funciones cognitivas como categorizar, percibir, recordar, inferir, deducir, generalizar, razonar, explicar, comprender y por ello se requiere su estudio en términos cognitivos, sobre todo si se tiene en cuenta que lo que se quiere enseñar es ciencia conceptual (Rodríguez y Moreira, 2002).

Para que el concepto entre a formar parte del bagaje de conocimientos del alumnado, se tiene que enlazar con los conocimientos previos que se poseen. Dentro de este aprendizaje conceptual se pueden distinguir a su vez dos tipos: a) Conocimiento conceptual primario que está siempre ligado a contextos; y b) Conocimiento conceptual reflexivo que ya no está ligado al contexto y supone que el alumnado ha sido capaz de realizar una abstracción (López, 2008).

Procedimientos

La enseñanza de las ciencias implica aprendizaje teórico y aprendizaje práctico. En la educación académica se parte de la idea de que la función de la instrucción es tanto la impartición del saber como del saber hacer, de tal manera que cuando el alumno concluya su formación, cuente con conocimientos y habilidades. En la realidad se encuentran muy vinculados los conocimientos sobre conceptos y procedimientos (Castañeda, 1998).

Los procedimientos son las habilidades, destrezas y estrategias cognitivas, manipulativas, comunicativas y de investigación de mayor o menor complejidad que los

alumnos deben saber utilizar para construir el conocimiento o para solucionar problemas (Cordón, 2008); los procedimientos “implican saber hacer algo, no sólo decirlo o comprenderlo” (Pozo, 2006, p.3).

Los procedimientos están asociados a las condiciones en las que se realiza la actividad, de manera que poseen un carácter más concreto ya que son una manera de actuar; es decir se trata de una serie de pasos o fases para conseguir determinado fin (Martínez y Bonachea, 2002); es la “forma de ejecutar alguna cosa” (Huerta, 2003, p.132).

Evaluación

La evaluación del aprendizaje es el conjunto de operaciones que tiene por finalidad valorar los logros alcanzados por los alumnos en el proceso de enseñanza aprendizaje con respecto a los objetivos planteados en los programas de estudio (Carreño, 2001).

La evaluación debe ser un proceso continuo de diagnóstico, monitoreo y comprobación final, del grado de eficiencia alcanzada en los resultados del trabajo desarrollado de modo continuo, constante, por maestros y alumnos durante el proceso de enseñanza aprendizaje, el cual parte de la determinación de los objetivos y concluye con un juicio de valor que expresa los resultados obtenidos en el logro de dichos objetivos (Delgado, 2006).

La evaluación debe ser formativa, es decir una evaluación de proceso, parte integral del proceso pedagógico. También debe ser cualitativa y constituir una observación analítica continua del proceso de aprendizaje para proporcionar retroalimentación permanente acerca de la eficacia de lo que se aprende y de la forma en que se aprende con el fin de permitir la modificación y perfeccionamiento de ambos (Jané, 2004).

Método

El presente trabajo se abordó desde la perspectiva teórica de la taxonomía de Bloom y de sus revisores Anderson (2004), Krathwohl (2002) y Mayer (2002), quienes conciben el proceso de aprendizaje desde dos dimensiones: la dimensión del conocimiento con sus categorías factual, conceptual, procedimental y metacognitiva y la dimensión del proceso cognitivo con sus seis categorías (recordar, comprender, aplicar, analizar, evaluar y crear). En este artículo, con respecto a la dimensión del conocimiento, sólo se abordan las categorías factual, conceptual y procedimental. La categoría metacognitiva no está contemplada dentro de los objetivos de esta investigación.

Como parte de los fundamentos teóricos del presente trabajo, se utilizaron principalmente los tipos de conceptos de relación y cuantitativos, que son los que se emplearon en la unidad estudiada.

Se utilizó el método de aprendizaje de conceptos por recepción (Moreira, 1999) y su transferencia (aplicación) por parte de los estudiantes porque se consideró que por ser la primera vez que el alumnado del 4º semestre del programa de Licenciatura en Economía tuvo contacto con este tipo de conceptos pertenecientes a una subrama específica de la economía agropecuaria, como lo es la economía de la empresa agraria, era necesario que el profesor presentara los conceptos y, a su vez, explicara y mostrara la aplicación de dichos conceptos para que sean recepcionados y posteriormente transferidos por los estudiantes en

diversos ejercicios planteados, además del reforzamiento de ese conocimiento mediante las lecturas correspondientes.

De igual manera, en el aprendizaje de procedimientos también se hizo uso del método de recepción y su posterior transferencia por parte de los alumnos por las mismas razones expuestas en el caso del aprendizaje de conceptos.

Para lograr el aprendizaje de los procedimientos, se hizo énfasis en la resolución de ejercicios, que fue uno de los componentes principales de esta investigación y en la interrelación que existe entre el aprendizaje de conceptos y el aprendizaje de procedimientos.

Las actividades de resolución de ejercicios se llevaron a cabo mediante el trabajo en grupos por parte de los discentes, con el propósito de que hubiera mayor nivel de aprendizaje por medio del trabajo colaborativo. A estos grupos se les llamó grupos permanentes de aprendizaje porque contaron con los mismos miembros durante toda la duración de la impartición de la Unidad V Economía de la Empresa Agraria.

En las dos secciones de estudiantes del 4º Semestre la Licenciatura en Economía (A y B), que estaban cursando la materia Economía Agropecuaria se formaron 8 grupos de aprendizaje (4 grupos por sección), siendo un total de 46 estudiantes, de los cuales 42 presentaron examen parcial.

Del total de los 8 grupos, 7 de ellos resolvieron 24 ejercicios cada uno durante seis sesiones en el aula. El equipo No. 4 de la sección A fue el único grupo que no efectuó dicha actividad, siendo además el último grupo de los ocho en formarse.

Al terminar el programa de la unidad, se impartió una sesión con el objetivo de despejar dudas o aclarar conceptos y procedimientos, y en la siguiente sesión se aplicó el examen parcial. Este último consistió en dos partes. La primera parte fue una prueba con reactivos de opción múltiple sobre los conceptos; la segunda parte fue una prueba que consistió en la resolución de ejercicios que exigió la utilización de los procedimientos correspondientes.

No se realizó la validación o prueba piloto del examen parcial porque éste fue elaborado tomando en consideración, por un lado, las definiciones conocidas de los conceptos estudiados; y por otro lado, los ejercicios a resolver fueron diseñados de manera muy similar a los resueltos por los grupos permanentes de aprendizaje.

Tanto en los ejercicios resueltos en el aula por los grupos permanentes de aprendizaje como en el examen parcial de la Unidad V Economía de la Empresa Agraria, se contemplaron los siguientes conceptos y procedimientos: a) Concepto y procedimiento de Beneficio; b) Concepto de Ecuación de Beneficio de la producción animal; c) Concepto y procedimiento de Margen bruto; d) Concepto y procedimiento de Utilidad Neta; e) Concepto y procedimiento de Flujo de Caja; f) Concepto y procedimiento de Margen de Mercadeo; g) Concepto y procedimiento de Punto de Equilibrio; h) Concepto de Ecuación de Punto de Equilibrio; i) Concepto y procedimiento de Presupuesto Equilibrado; j) Concepto y procedimiento de Presupuesto Parcial; k) Concepto de Ecuación de Presupuesto Parcial; l) Concepto y procedimiento de Maximización de Margen Bruto; m) Concepto y procedimiento de Análisis de Sensibilidad; n) Procedimiento de Análisis Interno; o) Concepto de Economía de la Empresa Agraria; y p) Concepto de Empresa Agraria.

Resultados

De acuerdo a la Tabla 1, los conceptos que fueron respondidos con mayor porcentaje de aciertos que otros, fueron Punto de Equilibrio, Beneficio, Ecuación del Beneficio de la Producción Animal, Margen Bruto, Margen de Mercadeo, Ecuación de Punto de Equilibrio, Maximización de Margen Bruto, Economía de la Empresa Agropecuaria y Empresa Agropecuaria, con porcentajes de respuesta correcta mayores al 80%. Los restantes conceptos estuvieron entre el 71 % y el 76 % de aciertos.

Poco más de la mitad de los conceptos fueron menos complicados de comprender, a diferencia de otros conceptos como Análisis de Sensibilidad, Presupuesto Parcial y la ecuación del Presupuesto Parcial que se dificultaron para la comprensión.

Tabla 1. Porcentajes de respuestas correctas e incorrectas a los conceptos

Concepto	Respuestas correctas	Respuestas incorrectas	Totales
Beneficio	88	12	100
Ecuación de beneficio de la producc. animal	88	12	100
Margen bruto	83	17	100
Utilidad neta	76	24	100
Flujo de caja	71	29	100
Margen de mercadeo	86	14	100
Punto de equilibrio	95	5	100
Ecuación de punto de equilibrio	93	7	100
Presupuesto equilibrado	74	26	100
Presupuesto parcial	69	31	100
Ecuación de presupuesto parcial	69	31	100
Maximización de margen bruto	81	19	100
Análisis de sensibilidad	64	36	100
Economía de la empresa agraria	88	12	100
Empresa agraria	86	14	100

Elaboración propia. Fuente: Trabajo de campo

En cuanto a los procedimientos, la Tabla 2 muestra que en el caso de los 11 procedimientos que fueron evaluados en el examen parcial, los relativos al Beneficio, Presupuesto Equilibrado y Maximización del Margen Bruto tuvieron un porcentaje absoluto de respuestas correctas (50% o más); en tanto que Flujo de Caja, Margen de Mercadeo, Análisis de Sensibilidad y Análisis Interno tuvieron un mayor porcentaje (sumado) de respuestas correctas y parcialmente correctas que respuestas incorrectas; y los procedimientos de Margen Bruto, Utilidad Neta, Punto de Equilibrio y Presupuesto Parcial tuvieron mayores porcentajes absolutos de respuestas incorrectas.

Las razones de todo esto se deben a que al momento de realizar el procedimiento, los alumnos omitieron algún paso de la secuencia, hicieron algún cálculo equivocado, agregaron variables de más, omitieron alguna variable, se confundieron con algún número, etc., es decir cometieron errores de defecto, de exceso o de sustitución.

En el caso del procedimiento de Presupuesto Equilibrado, todos los estudiantes lo llevaron a cabo correctamente. Esto significa que fue un procedimiento de alta elementalidad para los sustentantes.

Asimismo, con respecto al procedimiento de Presupuesto Parcial, éste fue el que mayor porcentaje de repuestas incorrectas tuvo, lo que permite inferir que para los alumnos fue altamente complicado comprenderlo bien a bien. Es decir no lograron relacionar correctamente las variables que contempla; lo mismo puede decirse con respecto a margen Bruto, Utilidad Neta, y Punto de Equilibrio, pues todos estos procedimientos se constituyen de diversas variables cuyas relaciones permiten obtener el resultado. No se demostró el dominio de los procedimientos y en estos casos se puede decir que no hubo un aprendizaje significativo.

Tabla 2. Porcentajes de respuestas correctas, parcialmente correctas e incorrectas a los procedimientos

Procedimiento	Respuestas correctas	Respuestas parcialmente correctas	Respuestas incorrectas	Totales
Beneficio	52	33	15	100
Margen bruto	12	24	64	100
Utilidad neta	12	24	64	100
Flujo de caja	45	31	24	100
Margen de mercadeo	19	43	38	100
Punto de equilibrio	33	12	55	100
Presupuesto equilibrado	100	0	0	100
Presupuesto parcial	7	17	76	100
Maximización del margen bruto	74	24	2	100
Análisis de sensibilidad	24	50	26	100
Análisis interno	48	45	7	100

Elaboración propia. Fuente: Trabajo de campo

Por otra parte, en la Tabla 3 se presenta la relación encontrada a nivel grupo permanente de aprendizaje (en las dos secciones) sobre el rendimiento en el aprendizaje conceptual y el rendimiento en el aprendizaje procedimental, expresados ambos en el porcentaje de aciertos en el examen parcial; en este caso considerando el conocimiento conceptual como simple y el conocimiento procedimental como complejo.

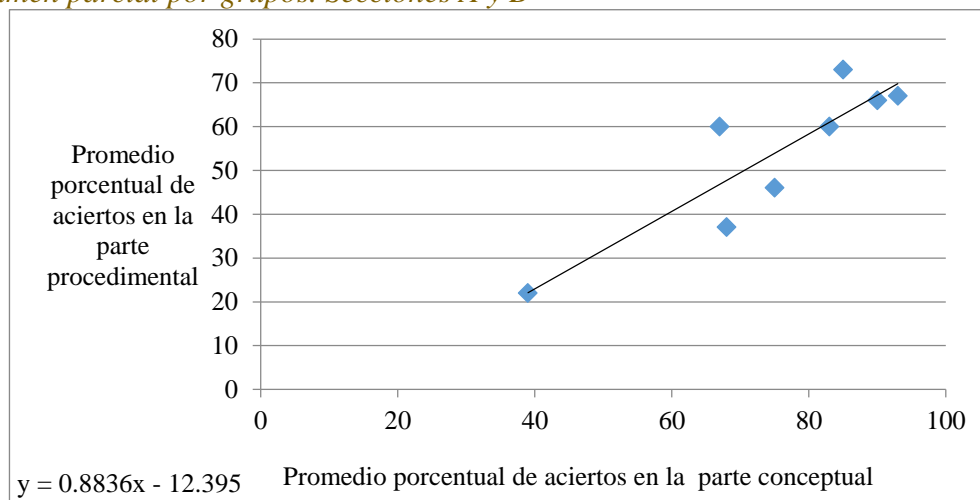
Tabla 3. Promedios porcentuales de aciertos en las partes conceptual y procedimental del examen parcial por grupos. Secciones A y B

Grupo y sección	Promedio porcentual de aciertos Conceptos.	Promedio porcentual de aciertos. Procedimientos.
1 A	68	37
2 A	93	67
3 A	90	66
4 A	39	22
1 B	85	73
2 B	75	46
3 B	83	60
4 B	67	60

Elaboración propia. Fuente: Trabajo de campo

En la Gráfica 1, la variable independiente (el aprendizaje conceptual medido en el porcentaje de aciertos en la parte conceptual del examen parcial) se ubica en el eje de las abscisas o eje de las X y la variable dependiente (el aprendizaje procedimental medido en el porcentaje de aciertos en la parte procedimental del mismo examen) se ubica en el eje de las ordenadas o eje de las Y.

Gráfica 1. Promedios porcentuales de aciertos en las partes conceptual y procedimental del examen parcial por grupos. Secciones A y B



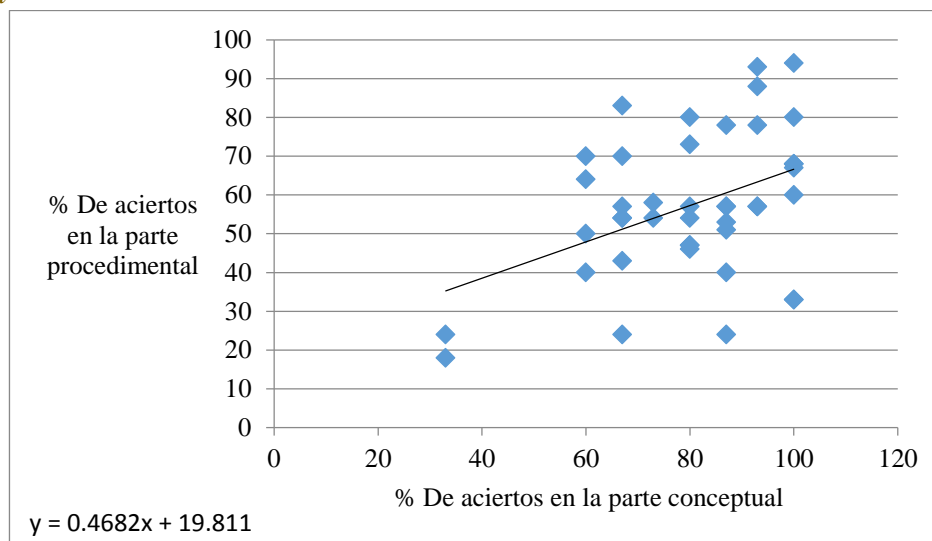
Elaboración propia. Fuente: Trabajo de campo

La Gráfica 1, que representa el comportamiento de cada uno de los 8 grupos, muestra en términos generales la tendencia de a mayor rendimiento en la parte conceptual del examen parcial, mayor rendimiento en la parte procedimental. La pendiente de la ecuación de la recta presenta signo positivo y por lo tanto determina una relación directa con respecto al comportamiento esperado. Se considera que del análisis de los cuadros y del diagrama de dispersión se desprende que en términos generales hay una tendencia que confirma el supuesto y por ende un logro de aprendizaje significativo en términos generales.

El hecho de que la mayoría de los estudiantes haya respondido correctamente a los conceptos y a la vez hayan respondido correctamente y parcialmente correctamente la mayoría de los respectivos procedimientos permite inferir que los educandos poseen ya una idea, al menos en términos generales, acerca de en qué consiste la Unidad y los diferentes componentes teóricos y prácticos que la integran, rebasando de esta manera el proceso cognitivo de recordar, mostrando el aprendizaje adquirido en los procesos de comprensión y aplicación.

En la Gráfica 2 (el diagrama de dispersión y la pendiente de la ecuación de la recta de la gráfica muestran una relación positiva entre ambas variables) se presenta en términos generales una tendencia que confirma en este trabajo la premisa de que al haber un mayor rendimiento en la parte conceptual del examen, entonces se presenta un mayor rendimiento en la parte procedimental del mismo, lo que permite asumir que además de que el conocimiento conceptual sirvió de base para adquirir conocimiento procedimental, esto se traduce también como un aprendizaje significativo para los estudiantes en cuanto a la aplicación de procedimientos, pues pudieron de esa manera comprender la estructura teórica de la economía de la empresa agraria y además aplicaron los conceptos que la conforman.

Gráfica 2. Porcentaje de aciertos en las partes conceptual y procedimental del examen parcial



Elaboración propia. Fuente: Trabajo de campo

Pero además es posible considerar que el aprendizaje fue significativo porque los alumnos rebasaron el aspecto puramente memorístico del aprendizaje al demostrar su manejo mediante la aplicación de la secuencias de los procedimientos correspondientes, lo que precisamente demuestran los datos que se han presentado.

El hecho de que hayan podido relacionar variables e ideas entre los conceptos y los procedimientos fue lo que le dio sentido, le dio significado a su aprendizaje y dio pie a la comprensión de la red conceptual estudiada en la Unidad en cuestión. Es decir, comprendieron que se pudo hacer algo con el aprendizaje: resolver ejercicios fundamentados en conceptos (que como parte de una teoría siempre son de relación) mediante la aplicación de algoritmos en los procedimientos.

Puede afirmarse que de esta manera se cumple con una de las premisas fundamentales de la taxonomía de Bloom y de sus revisores tales como Anderson y Krathwhol y Mayer, que es el establecimiento de un orden en la adquisición del aprendizaje, iniciando con los conocimientos simples hasta los conocimientos complejos.

En el caso de los resultados obtenidos en el examen parcial de la Unidad en cuestión, para cada grupo de aprendizaje de cada una de las 2 secciones de alumnos, se encontraron tendencias generales tanto al interior de cada grupo como de las 2 secciones, que permiten suponer que a medida que se adquieren conocimientos simples se está en capacidad de adquirir un conocimiento más complejo. Y esto se muestra relacionando los resultados de los aprovechamientos académicos de acuerdo a la siguiente relación de dependencia: en la medida en que el rendimiento en la prueba conceptual fue mayor, el rendimiento en la prueba procedimental también lo fue. Y esto da pie, entonces, a abordar esta situación desde la perspectiva del aprendizaje significativo.

Lo anterior concuerda con lo afirmado por Rivera (2004) de que un aprendizaje es significativo cuando el sujeto puede atribuir posibilidad de uso o utilidad al nuevo contenido, relacionándolo con el conocimiento previo.

Pero además, complementando todo lo anterior, la significatividad de los aprendizajes obtenidos en los conceptos (que en el caso de esta investigación fueron de relación y cuantitativos) y en los procedimientos en economía de la empresa agropecuaria, es precisamente porque dichos conceptos y sus correspondientes procedimientos forman parte de una estructura teórica o red conceptual. Esta cualidad es precisamente lo que consideramos que le dio sentido al aprendizaje y permitió pasar del conocimiento simple al conocimiento complejo.

En las Tablas 4 y 5 se presentan los tipos y cantidades de respuestas que los estudiantes de ambas secciones obtuvieron en el examen parcial, lo cual da una idea general acerca de su desempeño. En dichas tablas se muestra que en la parte conceptual del examen el 80% de las respuestas fue correcto, en tanto que en la parte procedimental el 64% de las respuestas son consideradas correctas o parcialmente correctas. Es decir, los estudiantes mostraron la adquisición de aprendizaje el cual, por la relación de dependencia establecida teóricamente (a mayor nivel de aprendizaje de conceptos, mayor nivel de aprendizaje de procedimientos) puede ser considerado como significativo.

Tabla 4. Tipo y cantidad de respuestas obtenidas en la parte conceptual del examen parcial

Tipo de respuesta	Cantidad	Porcentaje
Correcta	436	80
Incorrecta	110	20
Totales	546	100

Elaboración propia. Fuente: Trabajo de campo

Tabla 5. Tipo y cantidad de respuestas obtenidas en la parte procedimental del examen parcial

Tipo de respuesta	Cantidad	Porcentaje
Correcta	159	38
Parcialmente correcta	108	26
Incorrecta	153	36
Totales	420	100

Elaboración propia. Fuente: Trabajo de campo

Por todo lo anterior que se ha manejado en este trabajo, puede considerarse que los estudiantes, por los resultados académicos obtenidos en la Unidad y por las actividades que se efectuaron en la misma, en términos generales lograron un aprendizaje significativo tanto mediante la construcción de su conocimiento como por medio de los procesos cognitivos de la manera siguiente: los discentes, por una parte, construyeron su aprendizaje por medio de la lectura de los materiales acerca de los conceptos estudiados y por otra parte, mediante la práctica, a través de la aplicación de los algoritmos en la resolución de los ejercicios que se les asignaron por grupos en el aula. Actividades como éstas (la resolución de ejercicios con la supervisión del profesor) permiten que los sujetos puedan construir su propio conocimiento.

Y por otro lado, también se presentó un proceso cognitivo en el sentido de que los alumnos aprendieron conceptos y procedimientos que lograron asociar como parte de una estructura teórica que les permitió comprender las relaciones entre las variables que la constituyen y que le dio sentido a la adquisición del nuevo conocimiento. Por supuesto, los conocimientos previos que los educandos ya poseían desde antes de iniciar la Unidad,

formaron parte fundamental de ese proceso cognitivo que finalmente dio significado al aprendizaje de los contenidos de dicha Unidad.

Finalmente, en la sección A, el promedio de calificación obtenido por los alumnos fue de 63 puntos y en la sección B la calificación promedio obtenida por los estudiantes fue de 72 puntos.

En cuanto a las calificaciones obtenidas en las partes teórica y práctica del examen parcial se tiene lo siguiente: en la sección A se encontró que el total de los alumnos que formaron grupos permanentes de aprendizaje tuvieron un mejor desempeño en la parte teórica que en la parte práctica. En el caso de la parte teórica el rango de puntaje obtenido fue de 10 a 30 sobre 30 puntos, y en el caso del apartado práctico el rango de puntaje fue de 13 a 66 sobre 70.

En el caso de la sección B del total de alumnos que formaron los grupos también tuvieron un mejor desempeño en la parte teórica que en la parte práctica. En la parte teórica el puntaje tuvo un rango de 18 a 28 sobre 30, en tanto que en el apartado práctico el rango fue de 17 a 65 sobre 70 puntos.

Discusión

En términos generales, por un lado, se confirmó el supuesto teórico de que el aprendizaje menos complejo es el menos complicado de adquirir o construir, y esto quedó demostrado con los resultados de la investigación que indican precisamente que en la parte conceptual del examen se obtuvo un 80% de respuestas correctas y en el caso de la parte procedimental del mismo se logró un 64% de respuestas correctas y parcialmente correctas.

De hecho, en los conceptos que tuvieron su correspondiente procedimiento sólo en uno (Presupuesto equilibrado) se tuvieron más aciertos en la parte procedimental que en la parte conceptual por ser su procedimiento de alta elementalidad para los sujetos de este estudio.

En los restantes nueve conceptos que tuvieron sus respectivos procedimientos hubo mayor cantidad de respuestas correctas en la parte conceptual que en la parte procedimental.

Es necesario señalar que en el caso planteado en el párrafo anterior se cumple una de las afirmaciones de los revisores de la taxonomía de Bloom, con relación a que es posible que se presenten traslapes entre los diferentes niveles de aprendizaje en el sentido de que puede haber más dificultades de adquirir un aprendizaje aparentemente menos complicado que uno aparentemente más complicado. Así, para los discentes fue menos complicado efectuar el procedimiento de cálculo del Presupuesto equilibrado que definir algunos conceptos, como por ejemplo los conceptos de Presupuesto parcial y Análisis de sensibilidad y del propio Presupuesto equilibrado.

Por otra parte, los resultados que arrojó la investigación indican que a medida que se logró un mayor nivel de aprendizaje de conceptos, se logró un mayor nivel de aprendizaje de procedimientos, lo que implica el cumplimiento de esta premisa teórica, avalada no sólo por las relaciones de esas variables establecidas en las ecuaciones de regresión lineal sino también por los mismos datos de los resultados obtenidos en el examen parcial.

Así, en este caso un porcentaje significativo de procedimientos realizados en forma parcialmente correcta en el examen parcial permite inferir algún grado de dominio de los

mismos por parte de los alumnos, lo que implica en este caso cierto grado de aprendizaje sobre dichos procedimientos.

Referencias

- Amer, A. (2006). Reflections on Bloom's Revised taxonomy. *Electronic Journal of Research in Educational Psychology*, 8, vol 4 (1), 213-230.
- Anderson, L. W. (2004). Increasing teacher effectiveness. *International Institute for Educational Planning*. París: Organización de las Naciones Unidas para la Ciencia y la Cultura.
- Bloom, B. & colaboradores. (1986). Taxonomía de los objetivos de la educación. La clasificación de las metas educacionales. *Manuales, I y II*. Buenos Aires: El Ateneo.
- Carreño Huerta, F. (2001). *Enfoques y principios teóricos de la evaluación*. México, D.F.: Trillas.
- Castañeda Yáñez, M. (1998). *Análisis del aprendizaje de conceptos y procedimientos*. México, D.F.: Trillas.
- Cordón Aranda, R. (2008). *Enseñanza y aprendizaje de procedimientos científicos (contenidos procedimentales) en la educación secundaria obligatoria: análisis de la situación, dificultades y perspectivas*. Tesis presentada para obtener el grado de Doctor en el Programa de Investigación e Innovación en la enseñanza de las Ciencias. Murcia. Universidad de Murcia.
- Cruz Lara, X. & Anzaldo Cisneros, I. (2010). *¿Qué estrategias de aprendizaje utilizan los estudiantes de LIE? El caso de la UPN Unidad 142*. Veracruz: X Congreso Nacional de Investigación Educativa.
- Delgado Álvarez, Á. (2006). *Una estrategia de evaluación del aprendizaje en el nivel medio superior de la Universidad Autónoma de Guerrero: el ejemplo de la química orgánica en el 2º. Grado de la Unidad Académica No. 10*. Resumen de Tesis de doctorado en ciencias pedagógicas. La Habana: Instituto Central de Ciencias Pedagógicas.
- Huerta Ibarra, J. (2003). *Organización lógica de las experiencias de aprendizaje*. México, D.F.: Trillas.
- Jané, M. (2004). Evaluación del aprendizaje ¿Problema o herramienta? *Revista de Estudios Sociales*, 20. Bogotá: Facultad de Ciencias Sociales, Universidad de los Andes.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: an overview. *Theory into Practice*, 4, Vol. 4. College of Education, Ohio State University.
- López Sánchez, J. (2008). *Cómo aprendemos matemáticas para enseñarlas mejor*. Cádiz. Material didáctico del Centro de Enseñanza Infantil y Primaria San José Artesano.
- Martínez Verde, A. R. & Bonachea Montero, O. (2002). ¿Estrategia de enseñanza o estrategia de aprendizaje? *Revista Valera*, 3-6, Argentina.
- Mayer, Richard E. (2002). Rote versus Meaningful Learning. *Theory into Practice*, 4, Vol. 4. College of Education, Ohio State University.
- Moreira, M. A. (1999). *La teoría del aprendizaje significativo. Texto de apoyo No. 6*. Programa Internacional de Doctorado en Enseñanza de las Ciencias, Burgos: Universidad de Burgos & Universidade Federal do Rio Grande.
- Pozo, J. I. (2006). *Aprendices y maestros*. <http://www.slideshare.net/PATALEJ/capitulo-11-pozo>
- Rivera Muñoz, J. L. (2004). El aprendizaje significativo y la evaluación de los aprendizajes. Lima *Revista de Investigación Educativa*, 14 (8), Universidad Nacional Mayor de San Marcos.
- Rodríguez Nieto, M. C., Padilla Montemayor, V. M. & Esquivel Cruz, J. E.. (2007). Aproximación al aprendizaje. Punto de vista del estudiante. *Revista de Educación Superior*, 144, XXXVI., vol 4).
- Rodríguez Palmero, M. L. & Moreira, M. A. (2002). *Una aproximación cognitiva al aprendizaje del concepto "Célula": un estudio de caso*. Burgos. I Encuentro Iberoamericano sobre Investigación Básica en Educación en Ciencias.





Propiedades psicométricas de un instrumento para medir la dinámica familiar en niños

Psychometric Properties of a Scale Designed to Measure the Family Dynamics in Children

Esthela Jacqueline Madrid López, Angel Alberto Valdés Cuervo
Instituto Tecnológico de Sonora

María Dolores Valadez Sierra
Universidad de Guadalajara

Información adicional sobre este manuscrito escribir a:

Esthela Jacqueline Madrid López, esthelajacqueline.madridlopez@gmail.com, Angel Alberto Valdés Cuervo, angel.valdes@itson.edu.mx, María Dolores Valadez Sierra, doloresvaladez@yahoo.com.mx

Cómo citar este artículo:

Madrid López, E. J., Valdés Cuervo, A. A., & Valadez Sierra, M. D. (2015). Propiedades psicométricas de un instrumento para medir la dinámica familiar en niños. *Educación y Ciencia*, 4(43), 54-64.

Resumen

El estudio estableció las propiedades psicométricas de una escala para medir la dinámica familiar en niños. En el estudio participaron 664 estudiantes de escuelas públicas urbanas, de los cuales 386 (58.1%) fueron del sexo femenino (M edad=10.4, DE= 1.3 años) y 278 (41.9%) del masculino (M edad=10.5, DE=1.1 años). Los resultados señalan que la escala es empíricamente sustentable para medir el constructo mediante los factores de Autoridad Inductiva y Comunicación ($X^2 = 11.154$, $gl = 8$, $*p < .05$; $X^2/gl = 1.394$; CFI= .99; RMSEA= .024; GFI= .99). El instrumento presenta evidencias de validez concurrente ya que los puntajes de autoridad y comunicación se correlacionan de forma negativa con los reportes de victimización y agresión. Se concluyó que el instrumento puede ser utilizado satisfactoriamente en la evaluación de la dinámica familiar en niños.

Palabras clave: dinámica familiar, autoridad, comunicación, medición

Abstract

This study aims to establish the psychometric properties of a scale designed to measure the family dynamics in children. 664 students from urban public schools participated in the study of which 58% were female and 42% males.

Results sustain the internal validity of this instrument that aims to measure inductive authority and communication factors ($X^2 = 11.154$, $gl = 8$, $*p < .05$; $X^2/gl = 1.394$; CFI= .99; RMSEA= .024; GFI= .99). In addition, concurrent validity was supported by the negative correlations established between these factors and reports of victimization and aggression. It is concluded that the instrument can be used satisfactorily in the evaluation of family dynamics in children.

Keywords: standardization family dynamics, authority, communication, measurement

Introducción

La familia es el sistema donde se desarrollan relaciones afectivas más intensas que son la base del desarrollo y la socialización de los individuos. Los adultos, al realizar sus funciones como padres, influyen en la configuración de la conducta de los hijos (Aguirre, 2010). Este es el primer ambiente donde los individuos configuran sus estilos personales de interacción y adquieren las herramientas necesarias para establecer relaciones con otros individuos e integrarse en otros contextos sociales (Medellín, Rivera, López, Kanán & Rodríguez, 2012; Padilla & Díaz, 2011).

El contexto familiar constituye una de las principales influencias socializadoras de los individuos ya que en el mismo se adquieren habilidades, valores, expectativas y pautas de conducta que influyen en la relación con los otros contextos socializadores. Esto explica el hecho de que la familia posea un protagonismo especial en la prevención de conductas socialmente desadaptadas, como lo son los problemas de convivencia escolar (Smith, Schneider, Smith & Ananiadou, 2004; Yuste & Pérez, 2008).

El comportamiento social de los niños se asocia a características de la estructura y el funcionamiento de sus familias (Padilla & Díaz, 2011; Ruíz, Hernández, Mayrén & Vargas, 2014; Valdés, 2007). La afectividad de los padres es identificada como un aspecto clave de la crianza positiva vinculado con el bienestar socioemocional de los hijos (Kim-Cohen, Moffitt, Caspi & Taylor, 2004). Un entorno familiar bien estructurado y positivo, con relaciones cálidas, buena comunicación y una autoridad inductiva reduce los niveles generales de estrés en los niños y aumenta su probabilidad de lograr una adaptación positiva (Bowes, Maughan, Caspi, Moffitt & Arseneault, 2010, Coldwell, Pike & Dunn, 2006).

Para Friedemann y Smith (1997) las familias funcionales son capaces de crear un entorno que promueve el desarrollo cognitivo y socioemocional de sus miembros. La funcionalidad aparece asociada a características del sistema familiar tales como: cohesión, autoridad, comunicación, estrés, conflictos, emociones, vínculos entre otros (Olson, 1991).

La dinámica familiar se relaciona de forma específica con las prácticas de los padres dirigidas a crear un ambiente que permite a los hijos sentirse aceptados y valorados así como percibir que cuentan con su disponibilidad para apoyarlos en la solución de problemas que afrontan en sus vidas (Estévez, Martínez, Moreno, & Musitu 2006; Freeman & Newland, 2002; Maganto & Bartau, 2004; Sánchez & Valdés, 2011).

Los padres organizan y estructuran el sistema familiar a la vez que ejercen dos funciones básicas: comunicación y autoridad, ambas se interrelacionan y complementan mutuamente (Cataldo, 1991). El ejercicio de estas funciones exige modificaciones a través del tiempo, ajustándose en cada momento al crecimiento y desarrollo evolutivo de los miembros que componen el sistema familiar (Mestre, Samper, Tur & Díez, 2001).

La comunicación familiar es un aspecto clave en las relaciones familiares estrechamente vinculado con el ajuste psicosocial del adolescente (Musitu, Buelga, Lila & Cava, 2001; Martínez, Musitu, Murgui & Amador, 2009). La comunicación positiva entre padres e hijos, o en otros términos, una comunicación inductiva (es decir, abierta, fluida, respetuosa con los diferentes puntos de vista de los participantes y empática) favorece la aceptación social de los hijos y por tanto el ajuste escolar (Gaylord, Kitzmann & Lockwood, 2003; Steinberg & Morris, 2001).

La ausencia de control y apoyo familiar se vinculan con el desarrollo de conductas violentas en el ámbito escolar y con relaciones sociales problemáticas que favorecen el rechazo entre los iguales (Aguirre, 2010; Martínez, Musitu, Amador & Monreal, 2010). Según Darling (1999), las prácticas de los padres ejercen una influencia directa en las características de los jóvenes y se relacionan con la manifestación de síntomas conductuales tanto de naturaleza internalizante como externalizante en adolescentes (Hurtig, Taanila, Ebeling, Miettunen & Moilanen, 2005; Nishikawa, Sundbom & Hägglöf, 2010; Rothrauff, Cooney & An, 2009; Sandoval, Lemos & Vallejo, 2006).

La percepción de rechazo, las dificultades en la comunicación y las prácticas inadecuadas en el ejercicio de la autoridad son factores de riesgo que predicen la aparición de conductas agresivas y delictivas, ansiedad, depresión y quejas somáticas (García, Cerezo, De la Torre, Carpio, & Casanova, 2011). Del mismo modo, tanto la falta de supervisión como el excesivo control se relacionan con la frecuencia de aparición de comportamientos antisociales y disruptivos en población adolescente (Torrente & Vazsonyi, 2008).

Las familias de los niños con conductas agresivas se caracterizan por presentar alto nivel de conflicto parental, prácticas de crianza coercitivas o permisivas y brindar poco apoyo al hijo (Haynie et al., 2001). Los problemas de comunicación entre padres e hijos se encuentran asociados con el rechazo de los hijos por su grupo de iguales, con un pobre ajuste al contexto escolar y con una mayor participación en conductas violentas en la escuela, ya sea como agresores o como víctimas (Estévez, Musitu & Herrero, 2005; Franz & Gross, 2001; Gifford-Smith & Brownell, 2003).

Atendiendo a que la evidencia sugiere que la autoridad y la comunicación son aspectos de la dinámica familiar que afectan la conducta de los niños y la carencia de instrumentos validados en México para medir este constructo en estudiantes de educación básica, el estudio se propuso desarrollar un instrumento para medir este constructo en dicha población para lo cual se describen evidencias de validez (constructo y criterio) y confiabilidad de esta escala.

El estudio parte de la hipótesis de que la escala cuenta con evidencias de validez y confiabilidad para medir los aspectos de autoridad y comunicación de la dinámica familiar que le permiten utilizarse en la investigación de este constructo.

Método

Participantes

Para la administración de los instrumentos fueron seleccionados, de manera no probabilística, 664 estudiantes de escuelas primarias públicas urbanas de un municipio del noroeste de México. De estos, 386 (58.1%) fueron del sexo femenino (M edad=10.4, DE=1.3 años) y 278 (41.9%) del masculino (M edad=10.5, DE=1.1 años).

La muestra se dividió de forma aleatoria en dos submuestras de 332 estudiantes. La primera submuestra se utilizó para el análisis de confiabilidad y factorial exploratorio. Con la segunda se realizó el análisis confirmatorio del modelo de medida y se establecieron las evidencias de validez de criterio.

Instrumentos

Dinámica Familiar. Se realizó una adaptación del instrumento desarrollado por Sánchez y Valdés (2011) que midió el constructo a través de las escalas de Autoridad Inductiva (cuatro reactivos, ejemplo: mis padres me explican las razones de las reglas), donde se evalúa que los padres establecen y procuran el cumplimiento de las reglas explicadas a los hijos y Comunicación (cinco reactivos, ejemplo: siento confianza en mis padres para contarles mis problemas), que mide la vivencia del niño de contar con la ayuda y comprensión de los padres.

Se contestó mediante una escala tipo Likert con cinco opciones de respuesta que oscilaron desde 1 (Nunca), 2 (Casi nunca), 3 (A veces), 4 (Casi siempre) y 5 (Siempre).

Victimización por pares. Se utilizó el instrumento desarrollado por Valdés y Carlos (2014) para medir la frecuencia de reportes de victimización por parte de pares durante el último mes. Consta de 11 ítems que miden victimización (ejemplo: me esconden cosas, me pegan). Se contestó mediante una escala tipo Likert con cuatro opciones de respuesta, que fueron desde 1 (Nunca), 2 (Casi nunca 1-2 ocasiones), 3 (A veces 3-5 ocasiones), y 4 (Frecuentemente 5-10 ocasiones). Los índices de ajuste obtenidos en el análisis factorial confirmatorio ($X^2 = 27.23$, $p = .001$; $X^2/gl = 3.02$; CFI = .98; RMSEA = .055) demuestran sustentabilidad empírica del modelo de medición antes descrito (Blunch, 2013; Byrne, 2010).

Agresiones hacia los pares. Se utilizó el cuestionario de autoreporte desarrollado por Valdés y Carlos (2014) para medir la frecuencia de conductas de agresión durante el último mes. Este consta de seis ítems que miden la agresión (ejemplo: golpear a los compañeros, burlarse de los compañeros). Se contestó mediante una escala tipo Likert con cuatro opciones de respuesta, que fueron desde 1 (Nunca), 2 (Casi nunca 1-2 ocasiones), 3 (A veces 3-5 ocasiones), y 4 (Frecuentemente 5-10 ocasiones). Los índices de ajuste obtenidos en el análisis factorial confirmatorio ($X^2 = 19.73$, $p = .019$; $X^2/gl = 2.19$;

CFI = .99; RMSEA = .042) demuestran sustentabilidad empírica del modelo de medición antes descrito (Blunch, 2013; Byrne, 2010).

Procedimiento

Se explicó a las autoridades escolares el objetivo del estudio y se solicitó su autorización para acceder a las aulas. Asimismo, se obtuvo el consentimiento informado por escrito de los padres para que sus hijos participaran en el estudio. A los estudiantes se les invitó a participar de forma voluntaria garantizándoseles la confidencialidad de la información brindada por ellos.

En el análisis de las propiedades psicométricas del instrumento se establecieron evidencias de: (a) confiabilidad, determinada mediante la consistencia interna de los puntajes obtenidos mediante el Alfa de Cronbach, (b) validez de constructo, establecida mediante los análisis factoriales exploratorio y confirmatorio y (d) validez concurrente, mediante la relación de los puntajes con los reportes de agresión y victimización.

Resultados

Confiabilidad por reactivo

Se determinó la confiabilidad de los ítems mediante el estadístico Alfa de Cronbach. Se decidió mantener todos los ítems ya que sus correlaciones fueron superiores a .30 con relación a la escala (De Vellis, 2012; ver Tabla 1).

Tabla 1. Confiabilidad por ítems de la escala para medir dinámica familiar

Reactivos	Correlación del ítem-escala	Alfa eliminando el ítem	Decisión
Existen reglas en mi familia	.73	.92	Incluirlo
Mis padres me explican las razones de las reglas	.78	.92	Incluirlo
Mis padres hacen cumplir las reglas	.77	.92	Incluirlo
Las reglas de mi familia son adecuadas	.77	.92	Incluirlo
Puedo expresar y discutir con mis padres cuando no estoy de acuerdo con una regla	.47	.92	Incluirlo
Confío en mis padres para contarles mis problemas	.75	.92	Incluirlo
Mis padres me entienden	.77	.92	Incluirlo
Mis padres respetan mis intereses	.76	.92	Incluirlo
Estoy orgulloso de permanecer a mi familia	.79	.92	Incluirlo
Converso con mis padres acerca de mis problemas	.76	.92	Incluirlo

Validez de constructo

Análisis factorial exploratorio. Se utilizó el método de factores de ejes principales y rotación Oblimin. Se eliminaron los ítems con cargas factoriales menores a .30 y aquellos que presentaban cargas mayores a este valor en dos factores (De Vellis, 2012). Con base a este criterio se decidió no incluir el ítem "puedo expresar y discutir con mis padres cuando no estoy de acuerdo con una regla" de la escala.

Los valores del KMO de .93 y de la prueba de esfericidad de Bartlett ($X^2 = 4749.6$, $p < .000$) sugieren la adecuación de los datos para el análisis. Se obtuvo una solución de dos factores que explican el 71.37% de la varianza de los puntajes. La correlación entre los factores sostiene la idea de que ambos miden el mismo constructo, en este caso dinámica familiar (ver Tabla 2).

Tabla 2. Resultados del análisis factorial exploratorio de la escala para medir dinámica familiar

Ítem	Carga Factorial				
	<i>M</i>	<i>DE</i>	F1	F2	<i>h</i> ²
Existen reglas en mi familia	3.66	1.50	.00	.83	.69
Mis padres me explican las razones de las reglas	3.68	1.43	.01	.89	.77
Mis padres hacen cumplir las reglas	3.79	1.36	.03	.89	.75
Las reglas de mi familia son adecuadas	3.82	1.35	.16	.68	.67
Confío en mis padres para contarles mis problemas	3.78	1.44	.82	.00	.67
Mis padres me entienden	3.88	1.36	.91	.04	.76
Mis padres respetan mis intereses	3.85	1.41	.77	.06	.67
Estoy orgulloso de permanecer a mi familia	4.08	1.33	.64	.22	.68
Converso con mis padres acerca de mis problemas	3.79	1.45	.88	.04	.71
Correlaciones entre los factores					
Factor 1	-				
Factor 2	.78			-	

Análisis factorial confirmatorio. Se empleó el método de estimación de máxima verosimilitud para determinar la bondad de ajuste empírica del modelo. Se eliminaron tres ítems (Las reglas de mi familia son adecuadas, Confío en mis padres para contarles mis problemas y Estoy orgulloso de permanecer a mi familia) que afectaban el modelo ya que sus errores estaban relacionados (Byrne, 2010; Cea, 2004; ver Tabla 3).

Tabla 3. Solución Estandarizada por el Análisis Factorial Confirmatorio para un Modelo de Dos Factores

Ítem	Factores	
	Comunicación	Autoridad
Existen reglas en mi familia		.84**
Mis padres me explican las razones de las reglas		.90**
Mis padres hacen cumplir las reglas		.85**
Mis padres me entienden	.88**	
Mis padres respetan mis intereses	.85**	
Converso con mis padres acerca de mis problemas	.81**	

* $p < .05$, ** $p < .001$

Los índices de ajuste confirman la sustentabilidad empírica del modelo (Blunch, 2013; Byrne, 2010; Cea, 2004; ver Tabla 4).

Tabla 4. Índices de Ajuste del Modelo para Medir el Apoyo Familiar Afectivo

Modelo	gl	χ^2	χ^2/gl	CFI	RMSEA	GFI
Dos Factores	8	11.154*	1.394	.99	.02	.99

Nota. CFI=Índice de ajuste comparativo; RMSEA=Error de la raíz cuadrada media de aproximación;

GFI=Índice de bondad de ajuste

* $p < .05$. ** $p < .01$. *** $p < .001$.

Validez de criterio

Se estableció la evidencia de validez concurrente al establecer la correlación de los puntajes del instrumento con los reportes de agresión y victimización. Se encontró que los factores de comunicación y autoridad se relacionan de forma negativa con los reportes de agresión y victimización (ver Tabla 5).

Tabla 5. Medias, Desviaciones Estándar y Correlaciones entre Dinámica Familiar y Victimización

Variables	M	DS	1	2	3	4
1. Victimización por pares	1.65	.69	-			
2. Agresión hacia los pares	1.40	.57	.41***	-		
3. Comunicación	3.83	1.26	-.16***	-.23***	-	
4. Autoridad	3.70	1.30	-.15***	-.23***	-.68***	-

* $p < .05$. ** $p < .01$. *** $p < 0.001$.

Discusión

Los resultados sostienen la hipótesis de que la escala cuenta con evidencias de validez y confiabilidad que justifican su utilización en la investigación acerca de la dinámica familiar en niños mediante los aspectos relacionados con la autoridad inductiva y la comunicación.

El valor de la escala se refuerza por la solidez de los estudios con respecto a la influencia tanto de la autoridad como de la comunicación padre-hijo en el desarrollo infantil. El ejercicio de una autoridad inductiva por parte de los padres se asocia a un desarrollo infantil sano ya que protege al niño de situaciones de riesgo físico y psicológico, favorece la apropiación de las reglas sociales, favorece la autorregulación, el desarrollo moral y de conductas prosociales (Pereira, Canavarró, Cardoso & Mendoca, 2009; Roehlkepartain & Syvertsen, 2014). Por su parte la comunicación promueve que el niño se sienta apoyado, con la confianza para platicar con sus padres acerca de sus problemas y preocupaciones, lo que favorece la resiliencia a las dificultades de su etapa de desarrollo (Black & Logan, 1995; Sheridan, Stujs & Coutts, 2013).

La utilidad de la escala se fortalece por su validez concurrente que se evidencia en la correlación negativa significativa de sus puntajes con los reportes de victimización y agresión hacia los pares. Esto sugiere que la percepción de la dinámica familiar actúa como

factor protector de la violencia escolar ya que permite que las víctimas cuenten con el apoyo que necesitan para afrontar de forma efectiva la conducta agresiva de los pares (Lucy, Maughan, Caspi, Moffitt & Arseneault, 2010). Su relación negativa con los reportes de agresión es consistente con los reportes que evidencian que la calidad de la dinámica familiar favorece la estabilidad emocional y el desarrollo de emociones morales, que son aspectos que disminuyen las conductas agresivas en los niños (Kochenderfer-Ladd, Ladd & Kochel, 2009; Petrus & Van den Berg, 2013).

No obstante que los hallazgos sugieren que la escala es un instrumento valioso en la medición del constructo es necesario puntualizar que ésta posee limitaciones, ya que únicamente considera aspectos relativos a la relación padres-hijo para la evaluación de la dinámica familiar. Esto lleva a sugerir ampliar este instrumento con aspectos que evalúen la percepción de los hijos con respecto a la relación entre los padres.

Referencias

- Aguirre, A. (2010). *Prácticas de crianza y su relación con rasgos resilientes de niños y niñas*. Bogotá: Universidad Nacional de Colombia.
- Black, B., & Logan, A. (1995). Links between communication patterns in mother-child, father-child, and child-peer interactions and children social status. *Child Development*, 66, 255-271. <http://dx.doi.org/10.1111/j.1467.8624-1995.tb00869.x>
- Blunch, N. (2013). *Introduction to structural equation modeling using IBM SPSS Statistics and AMOS* (2nd ed.). Londres: SAGE.
- Bowes, L., Maughan, B., Caspi, A., Moffitt, T., & Arseneault, L. (2010). Families promote emotional and behavioural resilience to bullying: evidence of an environmental effect. *Journal of Child Psychology and Psychiatry*, 51, 809-817. <http://dx.doi.org/10.1111/j.1469-7610.2010.02216.x>
- Byrne, B. (2010). *Structural equation modeling with AMOS*. (2nd ed.). New York: Routledge Taylor & Francis Group.
- Cataldo, Z. (1991). *Aprendiendo a ser padres. Conceptos y contenidos para el diseño de programas de formación de padres*. Madrid: Visor.
- Cea, M. (2004). *Análisis multivariable. Teoría y práctica en la investigación social*. Madrid: Síntesis.
- Coldwell, J., Pike, A., & Dunn, J. (2006). Household chaos – links with parenting and child behaviour. *Journal of Child Psychology and Psychiatry*, 47, 1116–1122. <http://dx.doi.org/10.1111/j.1469-7610.2006.01655.x>
- De Vellis, R. (2012). *Scale development. Theory and applications*. New York: SAGE.
- Darling, N. (1999). *Parenting style and its correlates*. *Clearinghouse on elementary and early childhood education EDO-PS-99-3*. Recuperado de <http://www.athealth.com/Practitioner/ceduc/parentingstyles.html#Weiss>
- Estévez, E., Martínez, B., Moreno, D., & Musitu, G. (2006). Relaciones familiares, rechazo entre iguales y violencia escolar. *Cultura y Educación*, 18(3-4), 335-344.
- Estévez, E., Musitu, G., & Herrero, J. (2005). The influence of violent behavior and victimization at school on psychological distress: The role of parents and teachers. *Adolescence*, 40(157), 183-196.
- Franz, D. Z., & Gross, A. M. (2001). Child sociometric status and parent behaviors: an observational study. *Behavior Modification*, 25, 3-20. <http://dx.doi.org/10.1177/01454445501251001>
- Freeman, H., & Newland, L. (2002). Family transitions during the adolescent transition: Implications for parenting. *Adolescence*, 37(3), 457-475.
- Friedemann, M. L., & Smith, A. A. (1997). A triangulation approach to testing a family instrument. *Western Journal of Nursing Research*, 19, 364 -378. <http://dx.doi.org/10.1177/019394599701900305>
- García, M., Cerezo, M., De la Torre, M., Carpio, M., & Casanova P. (2011). Prácticas educativas paternas y problemas internalizantes y externalizantes en adolescentes españoles. *Psicothema*, 23(4), 654-659.

- Gaylord, N. K., Kitzmann, K. M., & Lockwood, R. L. (2003). Child characteristics as moderators of the association between family stress and children's externalizing, and peer rejection. *Journal of Child and Family Studies, 12*, 201-213. <http://dx.doi.org/1023/a:1022862816506>
- Gifford-Smith, M. E., & Brownell, C. A. (2003). Childhood peer relationships: social acceptance, friendships, and social network. *Journal of School Psychology, 41*, 235-284. [http://dx.doi.org/10.1016/S0022-4405\(03\)00048-7](http://dx.doi.org/10.1016/S0022-4405(03)00048-7)
- Haynie, D., Nansel, T., Eitel, P., Crump, A., Saylor, K., Yu, K. (2001). Bullies, victims, and bully/victims: Distinct groups of risk youth. *The Journal of Early Adolescence, 21(1)*, 29- 49. <http://dx.doi.org/10.1177/0272431601021001002>
- Hurtig, T., Taanila, A., Ebeling, H., Miettunen, J., & Moilanen, I. (2005) Attention and behavioural problems of Finnish adolescents may be related to family environment. *European Child & Adolescent Psychiatry, 14*, 471-478. <http://dx.doi.org/1007/s00787-0492-5>
- Kim-Cohen, J., Moffitt, T.E., Caspi, A., & Taylor, A. (2004). Genetic and environmental processes in young children's resilience and vulnerability to socioeconomic deprivation. *Child Development, 75*, 651-668. <http://dx.doi.org/10.1111/j.1467-8624.00699.x>
- Kochenderfer, B., Ladd, G., & Kochel, K. (2009). A child and environment framework for studying risk form peer victimization. En M Harris (Ed.). *Bullying, rejection, & peer victimization* (pp. 27-47). USA: Springer.
- Lucy, B., Maughan, B., Caspi, A., Moffitt, T., & Arseneault, L. (2010). Families promote emotional and behavioral resilience to bullying: evidence of an environmental effect. *Journal of Child Psychology and Psychiatry, 51*, 809-817. <http://dx.doi.org/10.1111/j.1469-7610.2010.02216.x>
- Maganto, J. M., & Bartau, I. (2004). *Corresponsabilidad familiar: fomentar la cooperación y responsabilidad de los hijos*. Madrid: Pirámide.
- Martínez, B., Musitu, G., Amador, L. V., & Monreal, M. C. (2010). Violencia escolar en adolescentes rechazados y aceptados: un análisis de sus relaciones con variables familiares y escolares. *Psicología: Teoría e Práctica, 12(2)*, 3-16. Recuperado de <http://www.redalyc.org/articulo.oa?id=193817420002>
- Martínez, B., Musitu, G., Murgui, S., & Amador, V. (2009). Conflicto marital, comunicación familiar y ajuste escolar en adolescentes. *Revista Mexicana de Psicología, 26(1)*, 27-40.
- Medellín, M., Rivera, M. E., López, J., Kanán, M. G., & Rodríguez, A. (2012). Funcionamiento familiar y su relación con las redes de apoyo social en una muestra de Morelia, México. *Salud Mental, 35(2)* 147-154. Recuperado de <http://www.redalyc.org/articulo.oa?id=58223312008>
- Mestre, M., Samper, P., Tur, A., & Díez, I. (2001). Estilos de crianza y desarrollo prosocial de los hijos. *Revista de Psicología General y Aplicada, 54(4)*, 691-703.
- Musitu, G., Buelga, S., Lila, M., & Cava, M. J. (2001). *Familia y adolescencia*. Madrid: Síntesis.
- Nishikawa, S., Sundbom, E., & Hägglöf, B. (2010). Influence of perceived parental rearing on adolescent self-concept and internalizing and externalizing problems in Japan. *Journal of Child and Family Studies, 19*, 57-66. <http://dx.doi.org/10.1007/s.10826-009-9821-y>
- Olson, D. H. (1991). Tipos de familia, estrés familiar y satisfacción con la familia: una perspectiva del ciclo vital. En C. J. Falicov (Ed.), *Transiciones de la vida familiar* (99-130). Buenos Aires: Amorrortu.
- Padilla, N., & Díaz, R. (2011). Funcionamiento familiar, locus de control y patrones de conducta sexual riesgosa en jóvenes universitarios. *Enseñanza e Investigación en Psicología, 16(2)* 309-322. Recuperado de <http://www.redalyc.org/articulo.oa?id=29222521007>
- Pereira, A., Canavarro, C., Cardoso, M., & Mendoza, D. (2009). Patterns of parental rearing styles and child behaviour problems among Portuguese school-aged children. *Journal of Family Studies, 18*, 454-464. <http://dx.doi.org/10.1007/s10826-9249-3>
- Petrus, A., & Van den Berg, E. (2013). Resilience in families in which a child is bullied. *British Journal of Guidance & Counselling, 41*, 504-517. <http://dx.doi.org/10.1080/03069885.2012.757692>
- Rochlkepartain, E., & Syvertsen, A. (2014). Family strengths and resilience: Insights from a National Study. *Reclaiming Children and Youth, 23(2)*, 13-18.
- Rothrauff, T., Cooney, T., & An, J. (2009). Remembered parenting styles and adjustment in middle and late adulthood. *Journal of Gerontology: Social Sciences, 64*, 137-146. <http://dx.doi.org/10.1093/geronb/gbn001>
- Ruíz, A., Hernández, M., Mayrén, P., & Vargas, M. (2014). Funcionamiento familiar de consumidores de sustancias adictivas con y sin conducta delictiva. *Liberabit. Revista de Psicología, 20(1)* 109-119. Recuperado de <http://www.redalyc.org/articulo.oa?id=68631260010>

- Salguero, V. (1993). La familia y el hogar: un ámbito para el desarrollo psicológico del infante. En R. Jiménez (Ed.), *Familia. Memoria del 1^{er}. Encuentro Nacional de Investigadores sobre Familia* (pp. 297-317). Tlaxcala: CUEF.
- Sandoval, M., Lemos, S., & Vallejo, G. (2006). Self-reported competences and problems in Spanish adolescents: A normative study of the YSR. *Psicothema, 18*(4), 804-809.
- Sánchez, P., & Valdés, A. (2011). Una aproximación a la relación entre rendimiento académico y la dinámica y estructura familiar en estudiantes de primaria. *Revista Intercontinental de Psicología y Educación, 13*(2), 177-196.
- Sheridan, S., Sjuts, T., & Coutts, M. (2013). Understanding and promoting the development of resilience in families. In S. Goldstein & R. Brooks (Eds.), *Handbook of resilience in children* (pp. 143-160). New York: SAGE.
- Smith, J. D., Schneider, H., Smith, P. K., & Ananiadou, K. (2004). The effectiveness of Whole-School Antibullying Programs: A Synthesis of Evaluation Research. *School Psychology Review, 33*(4), 547-560.
- Steinberg, L., & Morris, A. S. (2001). Adolescent development. *Annual Review of Psychology, 52*(1), 83-110. <http://dx.doi.org/10.1146/annurev.psych.52.1.83>
- Torrente, G., & Vazsonyi, A.T. (2008). Brief Report: The salience of the family in antisocial and delinquent behavior among spanish adolescents. *The Journal of Genetic Psychology, 169*, 187-197. <http://dx.doi.org/10.3200/GNTP.169.2.187-198>
- Valdés, A. (2007). *Familia y desarrollo. Estrategias de intervención en terapia familiar*. México: Manual Moderno.
- Valdés, A., & Carlos, E. (2014). Relación entre el autoconcepto social, el clima familiar y el clima escolar con el bullying en estudiantes de secundaria. *Avances en Psicología Latinoamericana, 32*, 447-457. <http://dx.doi.org/10.12804/apl32.03.2014.07>
- Yuste, N., & Pérez, M. (2008). Las cuestiones familiares como causa de la violencia escolar según los padres. *European Journal of Education and Psychology, 1*(2), 19-27.





Retos en la medición de la inteligencia en México: Lecciones aprendidas de la estandarización de las escalas de Wechsler

Challenges Measuring Intelligence in Mexico: Lessons Learned from the Standardization of the Wechsler Scales

Pedro Sánchez Escobedo

Universidad Autónoma de Yucatán, México.

Información adicional sobre este manuscrito escribir a:

Pedro Sánchez Escobedo, psanchez@correo.uady.mx

Como citar este artículo:

Sánchez Escobedo, P. (2015). Challenges Measuring Intelligence in Mexico: Lessons Learned from the Standardization of the Wechsler Scales. *Educación y Ciencia*, 4(43), 65-79.

Resumen

Para evaluar la inteligencia en México, tanto en ambientes clínicos como en el sistema educativo mexicano, se utilizan ampliamente las tres escalas de Wechsler, en niños desde los tres años de edad hasta adultos mayores. En este artículo se analizan los procesos de adaptación y estabilización de estas escalas en México, y se reflexiona acerca de las dificultades y retos que implicaron estos procesos con la finalidad de identificar elementos claves que nos permitan mejorar, en el futuro, el proceso de adaptación de escalas estandarizadas en México. Las escalas de Wechsler son los instrumentos más utilizados para identificar discapacidad de aprendizaje, discapacidad intelectual, talento y sobredotación; por lo que la discusión del proceso de estandarización de adaptación en México es importante tanto desde el punto de vista psicométrico como desde el punto de vista práctico. Derivada de la experiencia de estandarizar estas escalas en México, este artículo pretende aportar a la discusión sobre las ventajas y dificultades de la adaptación de pruebas diseñadas en los Estados Unidos en México. Examina de manera crítica los procedimientos de recolección de datos, análisis de los mismos y la pertinencia de su uso. En particular, se identifican asuntos clave tanto técnicos como metodológicos y las implicaciones culturales de la medición de la inteligencia México, en un intento de facilitar las futuras adaptaciones y estandarización de ese tipo de pruebas.

Palabras clave: inteligencia, pruebas psicológicas, estandarización

Abstract

In both educational and clinical settings, the most frequently used instruments to measure intelligence in Mexico are the three Wechsler Scales (for children ages 3 years old through older adults). Because the Wechsler scales are the most important tools in screening for intellectual disability, learning difficulties, and giftedness; critical points for the standardization in Mexico of the three scales are discussed, such as differences in translation, formats, and appropriateness of items, and idiosyncrasies of test takers in Mexico. In addition, empirical evidence to help select which test to use when norms for a given age overlap is provided to guide the users. Derived from the experience in participating in the standardization to these scales in Mexico, this article contributes to the discussion on the advantages and limitations of using American made tests in other cultures. Beyond the psychometric properties of these tests, this article focuses on the process of test adaptation and norm development by critically examining the procedures of data collection and data analysis. Finally, methodological, technical and cultural issues of measurement of intelligence in Mexico are examined in an attempt to facilitate future adaptations and standardizations of intelligence tests developed in the United States and used in other countries.

Keywords: intelligence, psychological testing, standardization

Introduction

Standardization of a particular test in another context, refers to transformations that involve an adjustment of means and/or standard deviations of either individuals or groups, or both (Fischer, 2004).

Describing the process of adapting intelligence tests for new cultural contexts is important because the methods and procedures of adaptation can affect the reliability of the scores obtained and their interpretation in a specific setting.

The Wechsler tests have been adapted in many countries, for example, in the United Kingdom the adaptation of the WMS-III took the form of a validity and comparability study between the American norms and the scores of a representative sample of the UK population. This process is acceptable since the same language and many cultural conditions are similar to the context where this test was originally developed. Thus investigators needed only to provide sufficient information to allow well-informed use of the USA norms in the UK. However, in other English speaking countries, such as India, cultural differences, test familiarity and educational background of respondents were considered as confounding factors for the test results. Hence, investigators in this case carried out changes to items in six of the ten subtests and reported new Reliability Coefficients, and percentiles ranges based on Indian normative data (<http://pearsonclinical.in/solutions/clinical-assessment-intervention/adult/memory/wms-iii-india/>).

Khaleefa (2006) compared the WISC-III adaptations for Japan and Sudan and reported high levels of reliability for both countries and the same factor structure. The only consistent difference was better performance in visuospatial tasks for Sudanese children, whereas Japanese children performed better in verbal tasks.

After revising the standardization of the Wechsler Intelligence Scale for Children, Third Edition (WISC-III) in sixteen countries, Georgas, Weiss, Van de Vijver, and Saklofske (2003) suggested that in spite of differences across cultures, there were more similarities than differences among the translated forms. They observed that in every country this test has shown the same factor structure and similar psychometric properties. The authors warn, however, that the interpretation and use of the test in each country should take into consideration the particularities of each culture.

The purpose of this paper is to reflect upon the process, results, and experiences of the adaptation and standardization of the three major Wechsler scales used to measure intelligence in educational, working, and clinical settings in Mexico. It is intended to identify advantages and limitations of the use of these scales, to pinpoint needs and challenges, and to facilitate future adaptations of new versions and standardizations or norms of these tests.

The Wechsler Scales in Mexico

Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III)

The Mexican Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III) is the most recently published test (2011). Its development was planned in two phases: the first was considered a national trial and comprised a sample of 1,801 children

from 21 of the 32 states in Mexico. This first attempt helped to identify the ambiguous translation of certain items and to correct some artifacts in the response protocols.

The actual standardization phase included 829 children from four major regions of the country, clustered in nine age related groups. Exclusion criteria included children with disabilities, obvious chronic disease and to those children whose birth language is not Spanish (for example, Maya, Nahuatl, etc.). Data for this article were collected from the national trial.

The WPPSI-III claims to place less emphasis on acquired knowledge than the other Wechsler tests and features shorter, more game-like activities that hold the attention of children as young as 2-1/2 years. Simplified instructions and scoring procedures enhance the ease of administration for examiners. Both children and examiners benefit from the thoughtful, carefully constructed revisions implemented to build a highly respected, reliable test that completely reflects what users wanted for preschool children. The range of testers is from 2 years 6 months to 7 year old.

Younger children take fewer subtests that aim to measure verbal comprehension and perceptual organization abilities. Older children take a greater number of subtests designed to measure verbal comprehension, perceptual organization, and processing speed abilities and the use of queries and prompts is generally unrestricted (<http://www.pearsonassessments.com>).

The Wechsler Intelligence Scale for Children – 4th Edition (WISC-IV)

The Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV) was published in 2003 in the United States and in 2007 in Mexico. This is the battery most widely used to assess IQ in both countries (Prifitera, Weiss, Saklofske, & Rolfhus, 2005; Sánchez-Escobedo, 2007).

The standardization sample for norm development consisted of 1,234 Mexican children in 11 age groups, with an average of 112 subjects per group. Participants were drawn from 12 of the 32 states in Mexico. Children with obvious physical or intellectual disabilities and those children whose first language is not Spanish were excluded from the sample. The sample was stratified by age, controlled for gender and type of school (private or public).

In Mexico, it was particularly important to standardize this test because the WISC-III had not been published or used in Mexico, and previously derived Mexican norms tended to overestimate Mexican children's IQ when compared with American norms (Padilla, Roll & Gomez Palacio, 1982; Esquivel, Heredia & Lucio, 2007).

Sánchez-Escobedo & Hollingworth (2012) examined the psychometric characteristics of this test, and through a confirmatory factor analysis and inter-correlational studies provided information on the WISC-IV. Factor loadings and correlational patterns were found to be comparable to those seen in the American versions of the test.

The Wechsler Adult Intelligence Scale - Third Edition (WAIS-III)

In 2001, a preliminary version of the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III) was adapted for use in Mexico to adjust norms solely for the group of reference. Then, a more extensive standardization process was carried out in 2003 to establish norms by age bands. In the preliminary phase, a translated version was

administered to a group of 287 persons between the ages of 16 and 70 years old controlling for gender and previous school experience, in an attempt to have a conventional sample, balanced for gender and educational level. In 2004, newly revised norms were published to adjust some discrepancies reported by users of this test.

Results from this first effort set the foundation for the standardization process, adapting and improving the previous version of the test and administering the battery to 970 Mexicans, from six different regions of the country. Previous experience helped to discard about 3% of tested individuals due to severe violations of the protocol of administration. The standardization sample controlled for gender, region of the country, and age. Due to sample size, norms were developed for 10 age bands, in contrast to the 14 bands in the American test.

In 2014, a new version of the test The WASIS-IV was published. This new version that claims to consider new demographic and clinical tendencies will gradually replace the older version of this scale for adults in Mexico.

Table 1 depicts and compares some of the major features of each of the Mexican Wechsler scales analyzed in this study. In this section, the target demographic for each of the Wechsler scales is described and a brief history of the scales' adaptation for use in Mexico is provided.

Table 1. Comparative between the tests

		WIPSSI-IV	WISC-IV	WAIS-III
Sample Size	Pilot	1,801	-	287
	Final	829	1,234	970
Year	Pilot	2009	-	2001
	Final	2011	2007	2003
# Subtests	Verbal	5	5	7
	Non verbal	5	4	7
Age	Bands	9	17	13
	Ranges	2:6 -7:03	6-16:11	16-≥70

General comments of the Mexican Wechsler scales

Translation and management of items

According to Hambleton (1996), “the term ‘adaptation’ rather than ‘translation’ was preferred by the test adaptation guidelines committee because the former term is broader and seemed to more accurately reflect the process of preparing a test or instrument for use in a second language or culture” (p. 5). Adaptation refers to the translation of items and the verification that the way it was translated is meaningful and familiar to the test taker. The process of test adaptation also involves ordering items according to the degree of difficulty to a specific population.

In the case of the Mexican scales, adaptation was first performed by professional translators; the majority of items were simply translated. Further adaptation was then implemented after analyzing responses in the standardization processes. For instance, extra care was taken to adapt questions for idiomatic expressions, to adapt to information relevant to the Mexican geography and history; Mexican children were asked who Benito Juárez was (a Mexican hero), instead of ‘Abraham Lincoln’. In other cases, additional positive and correct responses were considered. For instance, the item “*colony*” translated

as “*colonia*” in Spanish could be answered correct when the respondent asserted “perfume or fragrance”, instead of the intended translation of “settlement in a distant territory”.

However ranking of items was different than the original American versions since items were ordered according to the Facility index derived from the responses in the standardization samples. This decision was based considering the need of fairness in the discontinuation criteria pre-established for the tests. By ordering items accordingly to Mexican indexes, respondents are less likely to meet the discontinuation criteria due to the original test structure. Of course, scales such as *Digit Span*, *Coding*, *Letter-Number Sequence*, *Symbol Search*, and *Cancellation* remained unchanged.

Sampling methods

In Mexico, none of these tests used a randomly stratified sampling method due to many constraints: for example, budget restrictions, access issues in some school districts and regions, and the lack of an extended research network or acceptance of intelligence testing in the country. Hence, there is always the question about sampling methods and the degree to which the participants represent the intended population. More data should help assess this matter. For example, in every case, more than 50% of the sample size of the original American standardization procedures was reached.

For the standardization of the three tests, an effort was made to achieve some sort of balance of representation in the sampling. For example, there were attempts to balance for previous education, gender, and regional cultures of the country.

Norms derived from the standardization process of the Wechsler scales are representative of the population included in the sample, that is, for children and adolescents in the school system without obvious physical or mental disabilities. The adult scales are useful for functioning adults with various degrees of education and work experience. Rather than asking how representative the sample is for the Mexican population, the examiner must consider whether the subject to be tested resembles the characteristics of the Mexican standardization sample. Clear exclusion criteria were established for the three tests: they are not to be administered to people with obvious physical or mental disabilities, to people from rural communities, or to people whose native language is not Spanish. Thus, none of the tests have norms for special populations.

Test administrators

Test administrators in the United States are usually experienced psychologists or school counselors. In contrast, in Mexico advanced psychology students who were specifically trained for the purposes of the pilot testing administered most of the tests. Despite the training, results evidenced in some cases demonstrated a lack of care, a breach of pre-established protocols, or a lack of abidance to pre-established sampling criteria. In the future, testing conditions, scoring and data analysis should become a minimal source of error.

Differences in Record Forms and administration

Response protocols follow the general design of the original tests. The three Record Forms were reviewed and qualitatively compared. All three tests are completely written in Spanish.

In the case of the WISC-IV, on the Analysis page, the most salient difference is that the Mexican version uses a pre-established statistical significance level of $p \geq .05$ to estimate discrepancies and to facilitate scoring, since this is the common significance level used for interpretation. In the Mexican standardization process, consulted experts suggested that the inclusion of these figures would encourage screening for strengths and weaknesses. Likewise, the Mexican version uses larger fonts and figures than the American versions.

In the Mexican version, Word Reasoning, was removed for testing, norming, and reliability issues, because it made no sense to most Mexican children. For instance, questions such as “goes out at night” lead to responses like “my mom, when dad is on a trip”, when it was supposed to be a lead for “Moon”.

Another major difference was that in the case of Mexico, the manual allows for a break anywhere from 20 minutes up to 24 hours, to prevent fatigue of the children. This was derived from the observations during the standardization process that many of the Mexican children in the pilot group grew tired and distracted, due to their lack of previous exposure to testing routines such as this one. Formal recognition of the detrimental effects on the scores due to the lack of exposure and practice of these types of tests in Mexican children was an important advance.

Lessons Learned

Lessons learned led investigators to check and revise protocols to search for errors, inconsistencies, and even simulated responses in some rare cases. Over time, more rigid controls and supervision of data were necessary to avoid the problems reported in earlier phases of the test standardization process in Mexico. Data spreadsheets were audited up to three times to prevent missing or erroneous data, before analysis. Errors due to scoring decreased with every other standardization process beginning with the WAIS-III adaptation process in 1997, to the WIPSSI-III standardization in 2011.

In all three processes, the method of linear standard scores was used to calibrate norms and traditional statistical methods were used to analyze data, to correct errors, and to adjust norms. Scores in every test allowed for the establishment of scaled scores for the 10 core tests to calculate the four different indices: Verbal Comprehension, Perceptual Reasoning, Working Memory and the Processing Speed; and the total IQ. Table 2 depicts the difference among the three tests subscales.

Different statistical procedures were used to adjust means when atypical deviations from the expected mean of 10 was observed in a given sub-scale in a particular age band, attributing such deviation to sampling, scoring, or random errors. In future standardizations, investigators should ponder the effects of adjusting American norms with the Mexican derived data, rather than developing norms from scratch. For this, methods of mean and linear equating should be considered.

Table 2. Comparison of subscales across tests

Sub-tests	WPPSI-III		WISC-IV		WAIS-III	
	Type	# items	Type	# items	Type	# items
Block Design	E	0 - 40	E	0 - 68	E	0 - 68
Similarities	S	0 - 46	E	0 - 44	E	0 - 33
Information	E	0 - 34	S	0 - 33	E	0 - 28
Picture completion	S	0 - 32	S	0 - 38	E	0 - 25
Vocabulary	E	0 - 43	E	0 - 68	E	0 - 66
Matrix reasoning	E	0 - 29	E	0 - 35	E	0 - 26
Comprehension	S	0 - 38	E	0 - 42	E	0 - 33
Symbol search	s	0 - 50	E	0 - 60	E	0 - 60
Picture concepts	E	0 - 50	E	0 - 28	-	-
Coding	E	1 - 65	E	1 - 65	-	-
Letter-number	-	-	E	0 - 30	E	0 - 21
Digit span	-	-	E	0 - 32	E	0 - 30
Arithmetic	-	-	S	0 - 34	E	0 - 22
Word reasoning	E	0 - 28	S	0 - 24	-	-
Digit-symbol coding	-	-	-	-	E	0 - 133
Picture arrangement	-	-	-	-	E	0 - 22
Object assembly	-	-	-	-	E	0 - 52
Cancellation	-	-	S	0 - 136	-	-
Picture naming	S	0 - 30	-	-	-	-

Legend: S = supplementary; E = essential.

Which test to use?

One common question for practitioners is which test to use when two versions are available for the same age band. For instance, a seven-year-old child suspected of a learning disability can be tested with either the WIPSSI-III or the WISC-IV. In theory, the former test is easier and designed for younger children, and the upper bands of norms would be used to assess performance. The latter conveys more difficult demands, so the lower bands of norms are to be used. However the two instruments are different, containing different items and tasks. In the next section, we provide Mexican practitioners with the empirical information to guide decisions regarding which test to use.

From available data, raw scores were converted into standardized scores using norms derived for each test (WIPSSI, WISC or WAIS) and their overall mean calculated. Considering a theoretical mean of 10 for the population pre-established for each scale, differences in the means from the samples were inspected. Table 3 illustrates the ways overlapping age bands were used for this analysis.

Contrary to the general idea of this test being easier for older children, the use of the WIPSSI-III generally provided, in general lower scores, proving to be more difficult than the WISC-IV. Hence, when faced with the choice of the two tests, clinicians should be aware that children assessed with the WIPSSI-III will show lower scores, and thus will be more likely to be within ranges of mental disability and learning disorders, and less likely

to achieve standards for giftedness. Data would support advantages in the use of the WISC-IV when both tests are available.

Regarding the use of the WISC or the WAIS with 16-year-old children, there seems to be the same degree of bias in different directions: the WISC underestimates abilities whilst the WAIS overestimates them. Of course, further insight into which test to use will require administering both tests to the same subjects and compare the results. This proves to be an interesting venue for future research.

Table 3. Comparison of Mean of standardized scores in the sample per subscale in overlapping ages

Age 6	M-WIPSSI-III		DIF	M-WISC-IV		DIF
	M	SD		M	SD	
CD	7.88	(3.05)	-2.12	11.33	(3.61)	+1.33
VC	6.86	(2.95)	-3.14	12.50	(4.73)	+2.50
SM	6.64	(3.32)	-3.36	9.40	(4.97)	-0.60
MX	7.65	(2.48)	-2.35	11.16	(3.83)	+1.16
IF	7.45	(2.26)	-2.55	10.90	(4.14)	+0.90
SY	8.1	(3)	-1.90	10.1	(2.99)	+0.10
CM	6.87	(2.80)	-3.13	10.69	(3.91)	+0.69
Trend			-2.65			+0.86
Age 7	M-WIPSSI-III		DIF	M-WISC-IV		DIF
CD	8.27	(2.52)		-1.73	10.92	
VC	6.27	(3.07)	-3.73	11.15	(4.27)	+1.15
SM	7.53	(3.14)	-2.47	9.94	(4.54)	-0.06
MX	6.55	(2.47)	-3.45	10.55	(4.01)	+0.55
IF	6.63	(2.57)	-3.37	10.37	(4.17)	+0.37
SY	8.1	(3)	-1.90	10.1	(2.99)	+0.10
CM	8.27	(2.52)	-1.73	10.92	(4.32)	+0.92
Trend			-2.63			+0.56
Age 16	M-WISC-IV		DIF	M-WAIS-III		DIF
CD	9.07	(3.54)		-0.93	10.22	
VC	7	(3.98)	-3.00	16.69	(4.56)	+6.69
SM	8.59	(4.05)	-1.41	11.37	(0.65)	+1.37
MX	8.24	(3.28)	-1.76	10.20	(3.19)	+0.20
IF	9.07	(3.54)	-0.93	10.22	(3.20)	+0.22
SY	8.1	(4.3)	-1.90	10.1	(2.97)	+0.10
CM	8	(4.24)	-2.00	10.18	(3.02)	+0.18
Trend			-1.70			+1.28

DIFF: Difference from the theoretical mean of 10.

Additionally, Sanchez and Hollingworth (2010) have previously reported potential differences in score interpretation, depending on the norms used to create standardized profiles. The authors compared American, Hispanic, and Mexican norms in the WISC-IV, and reported that American norms tend to underestimate the IQ of high aptitude 7-year-olds, whereas they tend to overestimate the performance of low aptitude 16-year-olds. In almost every case, the Mexican norms tend to differ more from the American norms than Hispanic norms, and Mexican norms tend to overestimate Perceptual Reasoning and Verbal Comprehension when compared to American norms.

Although differences in standardized scores are expected when using different sets of norms, this continues to be a major issue regarding the external validity of the test.

Indeed, deciding the appropriateness of which set of norms or test is best to use needs further empirical data, derived from clinical and educational studies that depict how these batteries behave when confronting populations with known characteristics thoroughly documented by other clinical, psychometric, neurologic and image data. Table 3, compares mean standard scorers in overlapping ages.

Changes with age

By consider the three tests altogether, it was possible to explore how test adaptation behaves across the life span in the Mexican population. A comparison of variations in standardized scores across age was performed using 3 scales that persist across the life span. Figure 1 illustrates how the means of transformed scores for cubes design, vocabulary, comprehension and digit span, derived from the raw scores from the standardization samples, depart from the theoretical mean of 10 across the life span. It can be observed that there is a random variation from the expected theoretical mean of ten, either above or below the mean, with a mean difference of .6. Most likely, this minor variation is due to measurement error, sampling issues, or other testing artifacts, and they have little significance in practice, since variations seems to occur without a clearly identified trend.

However, clear underestimation of IQ above 60 years old is evident in this graphic. Thus, the future standardization of the next WAIS should pay attention to the development of norms for age bands above 70 years old, considering that currently, 9.7% of the Mexican population is older than 60 and the consequent need for better norms for age bands after 65 years old (UNAM, 2014).

Comparison by laterality and gender

Given that gender and laterality differences have been considered important in interpretation of scores, these variables were examined across the three tests. Table 4 compares means across the tests by gender. It can be seen that no significant gender differences were found across the standardization processes; this is in contrast with the expected slight superiority of females on verbal abilities reported in the extensive meta-analysis of Hyde and Marcia (1998).

Table 5 likewise shows that laterality was not a factor of significant statistical differences, except for vocabulary and matrix reasoning favoring left-handed people. Such differences may not be of practical significance and it could be assumed consistent with the arguments of Nettle (2003) arguing no practical differences in hand laterality and cognitive ability.

Table 4. Comparison of subscales by gender

Sub-tests	WPPSI-III			WISC-IV			WAIS-III		
	M	F		M	F		M	F	
Gender	X (SD)	X (SD)	t p	X (SD)	X (SD)	t p	X (SD)	X (SD)	t p
Block Design	23.06 (7.40)	22.85 (6.65)	.33 .73	31.39 (14.32)	29.67 (14.27)	2.11 .035	37.12 (26.88)	31.92 (12.79)	3.9 .001*
Similarities	15.75 (10.04)	15.88 (9.95)	-.15 .87	16.92 (9.46)	16.59 (9.35)	.62 .53	19.21 (12.27)	17.63 (6.68)	2.52 .012*
Vocabulary	16.58 (8.02)	16.92 (8.15)	-4.7 .63	30.66 (10.92)	29.75 (11.29)	1.44 .15	36.38 (13.05)	34.19 (12.84)	2.62 .009
Matrix reasoning	12.81 (5.40)	12.55 (4.97)	.56 .57	16.95 (6.04)	16.90 (6.26)	.11 .90	15.28 (10.77)	13.86 (9.17)	2.21 .027
Comprehension	14.23 (8.11)	14.59 (8.22)	-.51 .60	20.17 (7.64)	20.06 (7.41)	.26 .79	17.46 (7.21)	17.66 (7.34)	1.69 .09
Picture completion	17.18 (7.68)	16.97 (6.55)	.34 .72	22.22 (7.41)	21.85 (7.02)	.88 .37	18.02 (4.86)	17.45 (11.70)	.97 .32

Table 5. Comparison of subscales by Laterality

	WISC-IV			WAIS-III		
	L	R	t	L	R	t
Sub-tests	X (SD)	X (SD)	[p]	X (SD)	X (SD)	[p]
Block Design	33.32 (15.83)	30.51 (14.28)	1.26 .20	35.96 (12.99)	34.35 (21.22)	.53 .59
Similarities	18.46 (9.60)	16.63 (9.40)	1.25 .21	19.01 (6.09)	18.37 (9.93)	.45 .64
Vocabulary	33.67 (12.51)	30.03 11.08	2.10 .036	35.27 (12.98)	35.26 (12.93)	.005 .99
Matrix reasoning	18.37 (7.10)	16.86 (6.09)	1.57 .11	21.25 (31.27)	14.18 (7.02)	4.96 .001*
Comprehension	21.16 (8.54)	20.01 (7.48)	.98 .32	17.88 (7.10)	17.01 (7.28)	.83 .40
Symbol search	23.06 (8.23)	22.7 (8.77)	.20 .84	28.11 (10.46)	25.69 (11.16)	1.51 .13
Picture completion	23.37 (7.31)	21.91 (7.26)	1.28 .19	18.39 (4.35)	17.71 (9.35)	.51 .60

Discussion

Advantages of adapting the Wechsler Scales

During the translation and/or adaption of any test, it is important to make sure that the test is more understandable to the test takers, the directions are easy to comprehend, and the items are ordered on an appropriate scale of difficulty. In general, translation of verbal routines seems to be appropriate, and the differences between the American and Mexican versions regarding language competencies are minor. On executive routines, there was less influence from erroneous contextual and cultural factors, because many pictures were adapted to portray Mexican children or situations.

Disadvantages of adapting the Wechsler Scales

Given that Mexicans are usually less exposed to standardized tests it is usually wise to provide more time and expanded explanations before testing.

In addition, testing conditions should be carefully revised because in many cases testing was performed in a classroom, in a laboratory, or even outdoors. It is also important to provide additional practice exercises prior to some of the subtests and to make sure that subjects understand directions and procedures (Geisinger, 1994).

Testers should be also trained and supervised. Practice and mastery in handling materials, providing directions and interacting with the test taker are conditions that should not be taken for granted.

The use of appropriate norms

Maybe the most crucial question in adapting and standardizing tests pertains to the appropriateness of the newly developed norms to judge the tester's performance. Further research is needed to study the effects of using different norms for a given version of the test. For example, what happens when the American version is interpreted with Hispanic or Mexican norms? The Mexican form requires the American norms for special populations and suggests the comparison between sets of norms in case of doubt.

Different standard scores are naturally different depending on which test is used. Therefore, the key issue here is to decide, considering the characteristics of the test taker, which instrument would produce the most reliable information to compare against other data, whether it was psychometric or qualitative in nature, in order to contribute to a decision about the case. The fact that test results are only one element of the decision making equation is of paramount importance to bear in mind, not only to judge the case under study, but to assess the virtues and limitations of a given test used in the process.

It is not surprising to see differences in results using the different set of norms, since the norms were developed for use with different populations. Different scores may be associated to meaningful differences in the characteristics of the population taking the test. For example, to understand how culturally different American and Mexican public schools are from one another, consider this: of the 87% of Mexican children attending the public Mexican educational system, 53% of these children started their formal education in first grade, and 90% of them attend school on part time basis (Santibañez, Vernez, and Razquin, 2005; INEGI, 2009; INEE, 2009). Furthermore, for immigrants, it is difficult to measure the degree of acculturation and thus decide whether to test in English or Spanish, and use Mexican, American or Hispanic norms.

Conclusions

In the adaptation of the Wechsler scales to Mexico, translations from the original versions are in general adequate, and language adaptations to different semantic variations according to the target population are appropriate. In these scales, major differences were found related to the number of items, its order of presentation, and the replacement of some items for others. In general, all format variations were designed to increase understanding of directions, facilitate administration and scoring of the test. In sum, test adaptation may reduce biases due to cultural differences, practice, familiarity with standardized tests and other factors affecting Mexicans performance on the test and increase fairness in intelligence assessment.

Regarding norms, as expected, differences in standardized scores derive when different sets of norms are used. A same raw score may yield different standardized scores depending upon the set of norms used. Further research is needed to elicit information to help us decide which scale is better and which set of norms are more appropriate to use.

Future research needs to be conducted using the two overlapping age scales with the same person, and should be evaluated with other clinical, social, and educational criteria, thus indicating which of the two provides a better index of intellectual functioning under specific circumstances and in accordance with different decision making needs. Perhaps the estimation of an average score, derived from the use of the 3 set of norms will provide a better estimate of the person's intelligence. Indeed, the

longtime debate around cultural fairness of intelligence testing is revived by the results shown in this work.

The debate on limitations and boundaries of popular intelligence tests used abroad is revived when considering options between the psychometric strengths of these scales and the challenges in using them in different cultural contexts. As Garcia-Coll and Magnuson (1999) assert, “basic psychological and behavioral constructs might not mean in one culture what they mean in another” (p. 10).

The interest of scholars in the advantages and limitations of using adapted versions of American tests and the challenges of developing norms in other cultural contexts continues to be a source of disagreement. For example, the debate between Suen and Greenspan (2008) and Sánchez-Escobedo and Hollingworth (2009) regarding the use of either the Mexican or the American versions of Wechsler Adult Intelligence Scale (WAIS-III) in a death penalty case, renewed the discussion over what test to use and how to interpret the results when cross-cultural issues are present.

In sum, the use of intelligence tests, despite cultural and contextual influences, are necessary for practitioners in need of empirical evidence to support a number of high-stakes decisions; however, the results of these tests should be interpreted with caution. Test adaptation across cultures is a renewed field of interest in school psychology that offers various interesting lanes for future research.

Referencias

- American Educational Research Association, A. P. (199). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Child, D. (1977). *Psychology and Teacher* (2nd ed.). Holt, Reinhart & Winston. L td.: London.
- Esquivel, F., Heredia, C., & Lucio, E. (1999). . *Psicodiagnostico Clinico del niño* (2da ed.). Mexico DF: El Manual Moderno.
- Esquivel, F., Heredia, C., & Lucio, E. (1999). *Psicodiagnostico Clinico del niño* (2da ed.). Mexico DF: El Manual Moderno.
- Fina, A., Sánchez-Escobedo, P., & Hollinworth, L. (2012). Annotations on Mexico's WISC-IV: A validity study. *Applied Neuropsychology: Child* , 1, 6-17.
- Fischer, R. (2004). Standardization to account for croos-cultural response bias. *Journal of Cross Cultural Psychology* , 263-267.
- Flanagan , D. P., & Kaufman, A. S. (2004). *Essentials of WISC-IV assessment*. Hoboken, NJ: John Wiley & Sons.
- García Coll, C., & Magnuson, K. (1999). Cultural influences on child development: Are we ready for a paradigm shift? . In n. A. Maste (Ed.), *Cultural processes in child development: The Minnesota symposium of child psychology*. (Vol. 29, pp. 1-24).
- Geisinger, K. F. (1994). *Cross-cultural normative assessment: Translation and adaption issues influencing the normative interpretation of assessment instruments*. *Psychological Assessment*, 6(4), 303-312.
- Geisinger, K. F. (1994). *Cultural influences on child development: are we ready for a paradigm shift?*
- Georgas, J., Weiss, L., Van de Vijver, F., & Saklofske, D. (2003). *Culture and Children's intelligence: Cross cultrual analysisi of the WISC-III*. San Diego: Elsevier.
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton., P. F. Merenda, C. D, & Spielberger (Ed.), *Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Mahway, NJ: Lawrence Erlbaum Associates.
- Hyde, J. S., & Linn, M. C. (1998). Gender differences in verbal ability: A meta-analysis. *Psychoiologial Bulletin* , 104 (1), 53-69.
- Instituto Nacional de Evaluación Educativa INEE. (2009). Retrieved from <http://www.inee.edu.mx/>

- Instituto Nacional de Geografía, Estadística e Informática INEGI. (2009). Retrieved from <http://www.inegi.org.mx/inegi/default.aspx>
- International Testing Commission International guidelines for test use. (2001). *International Journal of Testing* , 7, 91-106.
- Khaleefa, O. (2006). Adaptation of the WISC-III in Sudan and Japan: A cross cultural study. *Arapsynet.e.journal* , 149-154.
- Nettle, D. (2003). Hand Laterality and cognitive ability: A multiple regression approach. *Brain and Cognition* , 17-22.
- Ogbu, J. U. (1994). From cultural differences to differences in a cultural frame of reference. In M. Greenfield, R. R. & Cocking (Ed.), *Cross-cultural roots of minority child development* (p. 365-391). Hillsdale: NJ: Lawrence Erlbaum Associates .
- Padilla, E. R., Roll, S., & Gómez, P. M. (1981). Ejecución del WISC-R en adolescentes Mexicanos. *Interamerican Journal of Psychology* , 16(2), 122-128.
- Prifitera, A., Weiss, L. G., Saklofske, D. H., & Rolfhus, E. (2005). The WISC-IV in the clinical assessment context. In D. H. Saklofske, G. L., & Weiss (Ed.), *WISC-IV clinical use and interpretation Scientist-practitioner perspectives* (pp. 33-71). San Diego CA: Academic Press.
- Psychological, C. (2003). *The WISC-IV Technical and interpretive manual*. San Antonio, TX: Harcourt Assessment.
- Reschly, D. J. (1981). Psychological Testing in Educational Classification and Placement. In *American Psychologist* (Vol. 36, pp. 1094-1102). Iowa State University.
- Sánchez-Escobedo, P. (2007). *Validación y normas para México de Escala Wechsler de Inteligencia para Niños IV*. . México: El Manual Moderno.
- Sanchez, P. (2001). Capítulo 6: Validación Preliminar para México. In D. Tulsy, & J. Zhu, *Manual Técnico del WAIS-III*. México DF: Manual Moderno.
- Sanchez, P., Canton, B., & Sevilla, D. (1999). *Compendio de Educación Especial*. México DF: Manual Moderno.
- Santibañez, L., Vernez, G., & Razquin, P. (2005). *Education in Mexico: Challenges and opportunities*. Santa Monica, CA: The Rand Corporation.
- Suen, H., & Greenspan, S. (2008). Linguistic sensitivity does not require one to use grossly deficient norms: Why U.S. norms should be used with the Mexican WAIS-III. In *in capital cases. Psychology in Intellectual and Developmental Disabilities. Official publication of Division 33, American Psychological Association*. <http://www.apa.org/divisions/div33/docs%5Cd33news-current.pdf>.
- Tulsy, D., & Zhu, J. (2001). *Escala Wechsler de Inteligencia para adultos-III*. Mexico DF: el manual Moderno.
- U.S. Department of Education, N. C. (2009). *The Condition of Education. (NCES Publication NO. 2009081)*. Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs2009/2009081_1.pdf. For summary of school expenditures see Table A-34-1 at <http://nces.ed.gov/programs/coe/2009/section4/table-tot-1.asp>
- UNAM. (2014). *Para el 2050 mas de la cuarta parte de la población sera vieja en México*. Unidad León: Escuela Nacional de Estudios Superiores.
- Weiss, L. (2003). Culture and children's intelligence: Cross-cultural analysis of the WISC-III . In J. Weiss, L. Van de Vijver, Saklofske, & D. (Ed.), *Georgas* (p. 50). United States: San Diego, CA: Academic Press.





Sobre la Técnica de Puntajes de Propensión (Propensity Score Matching) y sus usos en la investigación en Educación

Reviewing of Propensity Score Matching and its Uses in Educational Research

Claudia Patricia Ovalle Ramírez
Universidad de Antioquia, Colombia.

Información adicional sobre este manuscrito escribir a:
Claudia Ovalle Ramírez, claudia.ovalle@udea.edu.co

Cómo citar este artículo:

Ovalle Ramírez, C. (2015). Sobre la técnica de Puntajes de Propensión (Propensity Score Matching) y sus usos en la investigación en Educación. *Educación y Ciencia*, 4(43), 80-89.

Resumen

La técnica de Propensity Score Matching no es nueva, pero su uso no ha sido extenso en la investigación en Educación. En el presente artículo se argumenta que la utilización de esta técnica es conveniente en investigación en educación porque permite solucionar dificultades como el sesgo de selección y la imposibilidad de llevar a cabo experimentos por motivos éticos u otros. En ese sentido desarrollar diseños cuasi-experimentales ayudados con técnicas estadísticas como Propensity Score Matching (PSM) puede ser una alternativa viable que arroja resultados más precisos (al nivel de los hallados en situaciones experimentales) sobre efectos de interés.

Palabras clave: Propensity Score Matching, investigación, diseño cuasi-experimental, educación, causalidad

Abstract

Propensity Score Matching is not a recent statistical technique yet. It has not been used extensively in educational research. In the following paper it is stated that the technique would be worth of use as researchers may not always have the resources or ideal conditions in order to carry out experiments because of ethical and resource constraints. Quasi-experimental designs may be an alternative approach. By using PSM (Propensity Score Matching) it is possible to obtain more accurate results.

Keywords: Propensity Score Matching, research, quasi-experimental design, education, causality

Introducción

El Emparejamiento por Puntajes de Propensión o Propensity Score Matching (PSM) tiene el propósito de fortalecer los argumentos sobre la causalidad de las relaciones entre variables en estudios cuasi-experimentales. Esto debido a que el PSM reduce el sesgo de selección, es decir, el sesgo que se genera porque no se pueden asignar de forma aleatoria los individuos a una intervención o a un grupo control, para balancear todas aquellas características que pueden afectar la estimación del efecto.

Debido a que es un método *ex post-facto*, el PSM empareja los casos que recibieron una intervención (como si fueran parte de un grupo tratamiento experimental) con uno o más casos que no reciben dicha intervención. En este sentido, la diferencia con un experimento radica en que el grupo control puede ser “creado” con base a la información disponible (por ejemplo, un grupo control creado a partir de datos censales de personas que no son participantes). Incluso en diseños experimentales donde se contempla un grupo experimental y uno control aleatoriamente asignados, el PSM puede servir para balancear aún más las diferencias entre estos grupos permitiendo mejores estimaciones.

El emparejamiento de los casos se hace por medio de puntajes de propensión. Estos puntajes son probabilidades calculadas con base a características observables de los casos que reducen las diferencias entre estos con independencia de que hayan recibido o no el tratamiento, al igual que ocurre con la asignación aleatoria en los experimentos controlados. En palabras de Vinha (2006): “Los estimadores no paramétricos de emparejamiento (puntajes de propensión) son empleados frecuentemente en estudios de evaluación. La idea general de esta metodología consiste en determinar el impacto del tratamiento en los casos tratados usando información de observaciones que han recibido tratamiento y de observaciones similares sin tratamiento para construir un contra-factual sin tratamiento” (p.21).

Por su parte Heinrich, Maffioli y Vasquez (2010) indican que PSM “usa información de grupo de unidades que no participan en la intervención para identificar que le habría pasado a las unidades participantes en la ausencia de intervención. Por medio de la comparación de cómo los resultados difieren entre los participantes en comparación con no-participantes que son equivalentes en sus características observables, es posible estimar los efectos de la intervención promediando las diferencias entre participantes y sus casos de comparación emparejados” (p.6).

El PSM es una técnica cuyos resultados son bastante aproximados a los obtenidos en las situaciones experimentales reales, en términos del cálculo de los impactos (de una variable independiente sobre una dependiente) y de sus errores estándar, pero evitando las dificultades (éticas y de recursos) de una situación experimental real. Estos beneficios se han comprobado en estudios que comparan resultados obtenidos con medios experimentales con los de grupos de comparación no obtenidos aleatoriamente (i.e., estudios observacionales).

Por ejemplo, en el artículo de Dehejia y Wahba (1999) se reporta la evaluación del efecto de un programa de entrenamiento laboral sobre el salario de los participantes haciendo comparaciones entre diferentes grupos: grupo tratamiento y grupo control (ambos obtenidos de forma aleatoria), y de estos con muestras de tipo observacional de personas que no son participantes del programa. Se compara inicialmente el grupo tratamiento con el grupo control, ambos compuestos por participantes del programa - pero diferenciados en la época de reclutamiento-. Se encuentra que existe un efecto en términos de una ganancia salarial de los sujetos del grupo tratamiento equivalente a US \$1,794 en comparación con el grupo control tras su participación en el programa. Este

efecto fue a su vez comparado con el efecto estimado en muestras de personas que no son participantes del programa, y cuya información socio-demográfica y de ganancias salariales se reporta en encuestas nacionales. Para el subgrupo de participantes de la encuesta con sexo masculino el efecto estimado (si hubieran recibido el programa) llega a US \$1,473, mientras que este llega a los US \$1,774 dólares para la muestra total de encuestados. Los efectos hallados en las muestras observacionales (encuesta) se aproximan bastante a los del grupo experimental, por lo cual se puede concluir que con los datos apropiados (características socio-demográficas, variables de respuesta como el salario, entre otros) los efectos pueden ser estimados en diseños observacionales como lo haría un diseño experimental.

Usos del PSM en educación

Estimaciones de los efectos como las del ejemplo anterior se logran por medio de técnicas como PSM. Esta técnica puede ser ampliamente usada en la investigación en educación para evitar las dificultades que se presentan con el sesgo de selección y con el desarrollo de paradigmas de experimentación en este tipo de investigación. En el campo de la educación sería éticamente incorrecto asignar niños a una institución educativa de baja calidad para poder analizar su efecto en variables como el desempeño del estudiante, por lo cual existe un sesgo antes de realizar cualquier análisis. Este sesgo consiste en que los estudiantes pueden presentar una serie de características previas al ingreso a la institución las cuales pueden afectar sus resultados académicos como el nivel educativo de los padres, la zona de residencia y las condiciones económicas. Estas son covariables que afectan el rendimiento académico (i.e., el efecto del barrio, del capital cultural, etc.) pero que no son los efectos que interesan al investigador y que pueden sesgar los resultados que éste obtiene ya que se puede confundir su efecto con el de las variables de tratamiento (i.e., calidad del colegio).

En la asignación aleatoria propia del paradigma experimental las características de los individuos son balanceadas por medio de la asignación al azar de los casos entre el grupo experimental y el grupo control. Para evitar los sesgos, estos grupos no deben diferir en otras características a excepción de las producidas por la intervención. Una alternativa a la imposibilidad de asignar sujetos a un tratamiento (como escuela de baja calidad vs alta calidad; colegio de jornada completa vs. jornada parcial; institución educativa pública vs. privada) consiste en establecer que las distribuciones de características (covariables) de los individuos son equivalentes, para luego estimar el efecto o las diferencias entre estos sujetos, así no hagan parte de una situación experimental. De esta manera y sin necesidad de colocar estudiantes en situaciones experimentales éticamente inapropiadas, los resultados que un estudiante obtendría si estuviera en el grupo de tratamiento pueden ser estimados. Esto se logra por medio del emparejamiento de casos con características (covariables) similares sin importar si se trata de casos que reciben o no el tratamiento. Para esto se usan puntajes de propensión o probabilidades de pertenecer a un grupo dadas ciertas características de los sujetos. Este procedimiento es esencial en el PSM y se explicará más adelante.

El PSM ha sido usado con éxito para responder a preguntas sobre el efecto de programas sociales (i.e., Hsieh y Hamilton, 2014), y puede ser empleado también en el campo educativo. Con esta técnica se podrían resolver preguntas de investigación interesantes a partir de datos provenientes de bases nacionales educativas o con datos provenientes de estudios observacionales: ¿cómo sería el desempeño de estudiantes de instituciones de baja calidad educativa si se les colocara en instituciones de alta calidad?, ¿cómo sería el desempeño de estudiantes de instituciones públicas si se les

asignara a instituciones privadas? y ¿cómo se vería afectado el logro de los estudiantes de jornada parcial si se les transfiriera a jornada completa?

Aplicación de la técnica

Para aplicar la técnica de PSM el primer paso es calcular el PS (puntaje de propensión o “propensity score”) por medio de regresión logística en el caso de que la variable independiente sea de tipo bivariado (por ejemplo, pertenecer a una escuela de alta calidad o baja calidad; participar en una escuela de jornada completa o parcial; atender una institución educativa pública o privada). Es decir, se determinará la probabilidad de que los estudiantes pertenezcan a una institución educativa (variable independiente) y no a otra, dadas ciertas características y covariables. El modelo de probabilidad que subyace al PSM es:

$$\Pr(D=1|X_i) = \pi_i = f(\beta X_i)$$

Este es un modelo de probabilidad binomial (Pr) de pertenecer a un grupo tratamiento (por ejemplo, colegio de jornada completa) dadas unas características X (las covariables), y que es una función f de los predictores lineales βX_i , donde π_i es la probabilidad esperada del estudiante i. Los coeficientes β , estimados mediante máxima verosimilitud, capturan el efecto de las variables X sobre la probabilidad de asistir a un colegio de ciertas características (i.e., jornada completa). Lo que se va a determinar es, ¿cuál sería el logro escolar de los estudiantes de los colegios de jornada completa $Y | X, D = 1$, si pertenecieran a un colegio sin esta característica (por ejemplo, de media jornada) $Y | X, D = 0$?

La diferencia entre estos logros escolares (el de la jornada completa menos el de la media jornada) es el efecto de pertenecer a un colegio con la característica de interés para el investigador (por ejemplo, la jornada completa). Utilizando los coeficientes β , se construye la distribución de logros escolares para estudiantes de colegios de jornada completa $F(Y_1 | X, D=1)$ y media jornada ($Y_0 | X, D=0$), y se utiliza $F(Y | X, D=1)$ para aproximarse a $F(Y | X, D=0)$. En términos de Imbens y Angrist (1994) “debido a que no se observa el mismo individuo bajo las dos condiciones (tratamiento y sin tratamiento) se debe hacer comparaciones entre diferentes individuos y estimar los efectos promedio del tratamiento para poder acercarnos a la estimación del efecto. Es decir, se define el efecto causal en términos de resultados potenciales o contra-factuales” (p. 467).

El cálculo de los puntajes de propensión (que tienen un rango de 0 a 1 por ser una probabilidad) incluirá en la ecuación de regresión todas las covariables que el investigador determine como pertinentes. Las características observables que no se incluyan en el modelo de regresión pueden desmejorar la predicción realizada, por lo que se sugiere que se haga una revisión completa de la literatura en el tema para no dejar variables importantes por fuera del modelo, ya que un número adecuado de estas variables (Hsieh y Hamilton, 2014 incluyen hasta 12) permitirá obtener estimaciones más acertadas de los efectos (Tanaka y Kawakami, 2015).

Aunque para este control de covariables existen técnicas como ANCOVA, no se recomienda su uso si no existen grupos con asignación aleatoria de los cuales obtener la información, ya que ANCOVA no controla por todas las covariables y sus relaciones (Lane y Gibss, 2015).

Una vez obtenidos los puntajes de propensión, estos serán usados para hacer emparejamiento entre las instituciones educativas (una para el grupo tratamiento y otra para el control) que se asemejen en dichos puntajes. Para este emparejamiento, los software estadísticos (R o SPSS) permiten aplicar un procedimiento de emparejamiento exacto a toda la muestra seleccionada, es decir los colegios del grupo tratamiento

tendrán un par o colegio del grupo control con PS idéntico. Otras formas de emparejamiento incluyen: emparejamiento con la puntuación más cercana, con un caliper (número que indica una diferencia máxima entre los PS) o el emparejamiento con el uso de distancias (i.e., Mahalanobis). De esta manera, se podrá distinguir cuáles instituciones tienen influencia en el desempeño de los estudiantes comparadas con un grupo control de instituciones que se les han emparejado estadísticamente pero que no reciben el tratamiento. Es importante tener en cuenta que el procedimiento no se limita a instituciones educativas, sino que se pueden utilizar otras unidades de análisis como los estudiantes.

Después del proceso de emparejamiento de los casos se comprobará que estos no presentan sesgo en el emparejamiento, por ejemplo, graficando un histograma de la distribución de los puntajes de propensión y los casos del grupo tratamiento y control que les corresponden a cada uno de estos puntajes. También se verifica que haya reducción en términos del promedio y de las diferencias estándar (d de Cohen) en los puntajes de propensión y de las covariables (comparación pre y post emparejamiento), para establecer que el emparejamiento fue exitoso. Con la muestra emparejada se puede hacer el cálculo del efecto promedio del tratamiento, el cual se basa en las diferencias entre el promedio de los resultados de los individuos del grupo tratamiento y el promedio de los resultados del grupo control que se le empareja con base a las covariables.

Los casos que no se emparejan se desechan y aquellos que fueron emparejados se emplean en una nueva base de datos para hacer análisis estadísticos de interés para el investigador. Por ejemplo, se puede ingresar en esta nueva base una variable de resultado (o variable dependiente), como desempeño en matemáticas, para establecer si las escuelas de jornada completa (grupo tratamiento) tienen mejor rendimiento comparadas con las escuelas emparejadas del grupo control (escuelas de media jornada). Esto se puede hacer aplicando de nuevo un modelo de regresión (el primero permitió el cálculo de los puntajes de propensión explicado anteriormente). El segundo análisis de regresión incluirá la variable dependiente (los puntajes en matemáticas), la variable independiente jornada (completa o media), junto con todas las covariables que sirvieron para efectos del emparejamiento: tamaño de la escuela (covariable1), porcentaje de estudiantes de minoría (covariable2), precio de la matrícula (covariable3), etc.

De acuerdo con la WWC (What Works Clearinghouse, 2012), en un cuasi-experimento, como aquellos en los que se usan técnicas como PSM, debe existir mínimo dos grupos diferenciados: el grupo que recibe la intervención (por ejemplo, instituciones de jornada completa, de tipo público o de alta calidad) y el grupo control (instituciones de media jornada, o de tipo privado o de baja calidad). Debido a que los grupos pueden diferir, se deben igualar en las características observables lo cual puede ser una fuente de sesgo. En la primera fase del análisis estadístico el criterio de similitud entre el grupo tratamiento y el grupo control consiste en que la diferencia de las características entre los mismos debe ser menor de 0.25 de una desviación estándar, por ejemplo al comparar las covariables (Ho, Inía, King, Stuart, 2007 en WWC, 2012).

Otros estándares que deben cumplir los diseños cuasi-experimentales son: inclusión de una variable dependiente válida y confiable (por ejemplo, una medida estandarizada en la población como los exámenes de último año de Bachillerato); existencia de más de una unidad de análisis en cada grupo (se sugiere un mínimo de 500 unidades en el grupo tratamiento), y garantía de que los efectos se atribuyen a la intervención de acuerdo con la literatura empírica en el área. Debido a las debilidades que en términos de su capacidad de inferencia tienen los estudios cuasi-experimentales, Sadish, Cook y Campbell (2002) recomiendan que en la estimación de efectos se

consideren elementos del diseño como la inclusión de puntuaciones pre-test, el emparejamiento con varios métodos para comparar sus resultados, y hacer uso de múltiples grupos de control.

Ejemplo empírico

Ten Brummelhuis y Vander Lippe (2010, citados por Lane 2015) presentan un ejemplo de la aplicación de la técnica de PSM en el área de RRHH (Recursos Humanos). Los autores utilizan como variable independiente la adopción de horarios flexibles de trabajo (FWA) y como dependiente la productividad y los comportamientos de ayuda a otros por parte de los empleados. Ya que no todo empleado estará satisfecho con este tipo de arreglo flexible se hace necesario identificar las características de quienes obtienen más beneficio con este estilo de trabajo (i.e, trabajadores casados), y así poder emitir una conclusión válida con respecto al efecto de este tipo de programas. Por tanto, se usan las covariables género, edad, educación, tareas del hogar, motivación del empleado, entre otras (tabla 1). Estas son características que se pueden relacionan con el hecho de que una persona asuma o no un horario flexible. Para una muestra de 482 trabajadores, se logra hacer un emparejamiento de 344 casos con PSM, y se encuentran las diferencias antes y después del emparejamiento.

Como se puede observar en la tabla 1, las diferencias entre los grupos de trabajadores con horarios flexibles y no flexibles se disminuyen después del emparejamiento para variables como autonomía y horas laborales. Estas variables parecían tener un efecto mayor en la comparación entre grupos antes del emparejamiento pero sus valores *d* (diferencia estandarizada) disminuyen después del emparejamiento (i.e., autonomía desciende de 0.67 a 0,08). Así mismo, la edad y el género parecen ganar en importancia en la selección de un horario de trabajo flexible después del emparejamiento ya que sus valores *d* aumentan (i.e., el efecto de la edad se convierte en positivo e incrementa de -0.3 a 0.04).

Tabla 1. Medias, desviaciones y diferencias medias estandarizadas (d) en las covariantes y puntajes de propensión antes y después del emparejamiento

Variable	Antes del emparejamiento				Después del emparejamiento			
	Media FWA	Medias no FWA	Desviaciones no FWA	d	Media FWA	Medias no FWA	Desviaciones no FWA	d
Puntaje Propensión	0.69	0.32	0.25	1.54	0.69	0.69	0.23	0.04
Mujer	0.46	0.71	0.47	-0.5	0.46	0.37	0.48	0.19
Edad	37.3	41.5	10.5	-0.3	37.3	36.8	12.54	0.04
Autonomía	3.92	3.39	0.74	0.67	3.92	3.85	0.74	0.08
Tareas de casa	9.94	13.45	7.88	-0.4	9.95	10.15	8.09	-0.0
Horas laborales	38.5	29.2	11.21	0.90	38.5	41.80	13.03	-0.3
Presión laboral	3.08	3.17	0.57	-0.1	3.08	3.18	0.64	-0.1

Nota: FWA es horario laboral flexible. Tomado de Ten Brummelhuis & Van Der Lippe, 2010.

Para estimar los efectos del horario laboral flexible, los 344 casos emparejados y sus puntuaciones se usan en un modelo de regresión que predice los comportamientos de ayuda del empleado. En este modelo la varianza total explicada de los comportamientos de ayuda fue de 4.7% que es mayor que la encontrada antes del emparejamiento (con una varianza total explicada del 3%). Por tanto, los autores concluyen a partir de la muestra emparejada que existe un efecto de los horarios flexibles en las conductas de ayuda de los empleados. Un análisis de los coeficientes permite determinar que este efecto se da principalmente para las personas solteras (la

variable tiene un peso $B = 0.16$, el mayor entre las comunalidades encontradas). Este efecto de la variable estado civil soltero no era detectado antes del emparejamiento (coeficiente Beta = -0.39). (Ten Brummelhuis & Van Der Lippe, 2010).

Software para los cálculos con PSM

Para llevar a cabo PSM se pueden usar varios paquetes estadísticos (R, SPSS, etc.). Randolph, Falbe, Kureethara y Balloun (2014) indican que para hacer un PSM con R, primero se debe instalar la versión gratuita de software en el sitio web de “R Project for Statistical Computing” (<https://www.r-project.org/>). A continuación se requiere instalar y cargar el paquete MatchIT (el cual se encuentra en el menú “Packages”). El paquete Match IT se descarga solo una vez pero se requiere cargarlo cada vez que se haga un análisis estadístico. Para correr el análisis se requiere una variable de agrupación (VI o tratamiento) y una o más variables de emparejamiento o covariables que permiten igualar a los grupos. Un ejemplo de variable tratamiento puede ser la pertenencia a una escuela pública (por oposición a la escuela privada), a la jornada completa (por oposición a la media jornada), o la asignación a un programa o tratamiento (i.e., un programa de prevención del uso de sustancias, un programa de mejoramiento de las habilidades académicas, etc.) que por lo general son variables dummy que se codifican como 1 o 0.

Los datos que se suben a R pueden venir en cualquier formato pero se recomienda que se guarden como archivos .csv (por ejemplo desde Excell). Este archivo debe incluir casos como filas y los valores de cada variable de interés en las columnas (i.e., las covariables y la variable tratamiento). La variable independiente es la variable de agrupación que indica si un caso pertenece a un tratamiento o a un grupo control y las covariables también deben incluirse en el archivo. El código en R para ingresar la información es el siguiente (tomado de Randolph et al., 2014):

```
mydata <- read.csv ("C:/r/newyork.csv")
attach (mydata)
mydata[1:10,]
```

Este código puede ser copiado reemplazando la información entre paréntesis de la primera línea por la ubicación del archivo a analizar (‘C’ es el drive donde se almacena la información; ‘r’ es la carpeta y “newyork” el nombre del archivo, que deben ser reemplazados por los que va a usar el investigador). A continuación se desarrolla el emparejamiento y la evaluación de los resultados con el siguiente código al cual se le deben reemplazar las variables incluidas por las que interesen al investigador:

```
m.out = matchit (stw ~ tot + min + dis, data = mydata, method = "nearest", ratio = 1)
summary (m.out)
Plot(m.out, type = "jitter")
Plot(m.out, type = "hist")
```

La primera línea hace el emparejamiento con el comando Matchit, siendo la variable independiente stw que en el ejemplo original de Randolph et al, 2014 se refiere a tener la condición de ser una escuela del programa Schools to Watch (stw). Las covariables son: tot, min, dis, donde “tot” hace referencia al tamaño de la escuela, “min” se refiere al porcentaje de estudiantes de minoría y “dis” se refiere al porcentaje de estudiantes que reciben almuerzo gratis o de precio reducido. Estas variables deben ser reemplazadas por la variable independiente y covariables que el investigador desee emplear en su análisis.

El comando ‘ratio’ indica que se emparejará uno a uno, es decir una institución educativa del grupo tratamiento con una del grupo control. Existen varios métodos de

emparejamiento: vecino más cercano (arroja las diferencias promedio más bajas entre los grupos), emparejamiento exacto (empareja cada unidad del grupo tratamiento con una unidad del grupo control que tenga exactamente los mismos valores en cada covariable), sub-clasificación (divide el set de datos en subgrupos cuya distribución de covariables es similar en cada subclase); emparejamiento óptimo (minimiza la distancia promedio entre todos los pares emparejados), y emparejamiento genético (emparejamiento con un algoritmo especial).

Los resultados del emparejamiento son guardados en una variable denominada `m.out`. Las líneas 2 y 3 del código (que inician con el comando `plot`) imprimen y hacen gráficos de esta variable. Los resultados del emparejamiento arrojan las medias para el grupo tratamiento y control, la diferencia promedio entre los grupos (antes y después del emparejamiento) en todas las variables y covariables incluidas en el modelo, y la desviación estándar con respecto al grupo control. Estos estadísticos deben mostrar disminución en su cantidad después del emparejamiento, de lo contrario este no es exitoso. En el reporte de resultados el software R también se incluyen las diferencias entre cuartiles entre el grupo control y tratamiento para todas las variables en términos de media, mediana y diferencia máxima entre cuartiles (entre menor sea este último número es mejor el emparejamiento).

Una vez que los datos están emparejados se vuelve a crear una base de datos que solo contiene estos casos emparejados. Esto se hace con el código siguiente (tomado de Randolph, et al., 2014):

```
m.data1 <- match.data(m.out)
write.csv(m.data1, file = "C:/r/newyork_nearest100.csv")
```

La primera línea crea un set de datos en R con los datos emparejados eliminando aquellos que no lo están, la segunda línea convierte el set emparejado en un nuevo archivo de tipo `csv`. En esta línea se debe reemplazar ‘newyork-nearest100’ por el nombre que se le quiere dar al nuevo archivo. También reemplazar ‘r’ por el archivo donde se quiere almacenar esta base. Con esta nueva base se puede hacer cualquier análisis estadístico que interese al investigador. Por ejemplo, se puede determinar si las escuelas de jornada completa (grupo tratamiento) obtienen mejores resultados que las escuelas de media jornada (grupo control) en pruebas estandarizadas de rendimiento, por medio de una nueva regresión que incluya todas las covariables.

Conclusiones

PSM es una técnica que permite obtener una estimación más ajustada, debido a la inclusión de covariables y a la disminución del sesgo de selección, cuando un efecto no se puede calcular con base a un diseño experimental. Es una técnica que hace uso del cálculo de puntajes de propensión o probabilidades de pertenencia a un grupo para establecer el emparejamiento entre casos de grupo tratamiento y un grupo comparación establecidos por medios estadísticos. Esta técnica no implica la manipulación de variables de forma experimental por parte del investigador lo cual puede ser una ventaja en investigación educativa. El software R permite a los investigadores en educación establecer muestras de casos de tratamiento y control emparejadas con base en puntajes de propensión en particular cuando trabajan con datos como los de bases de resultados de pruebas estandarizadas educativas.

Referencias

- Agodini, R. Dinaski, M. (2001). *Are experiments the only option? A look at dropout prevention programs*. Mathematica Policy Research Inc.
- Dehejia, R.H., Wahba, S. (1999). Causal effects in non-experimental studies: reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94 (448), 1053-1063.
- Heinrich, C., Maffioli, A., y Vasquez, G. (2010). *A primer for applying propensity score matching. Impact Evaluation Guidelines. Technical Notes N IDB-TN-161*. Interamerican Development Bank.
- Holmes, W.M (2014). *Using Propensity Scores in Quasi Experimental Designs*. Los Angeles: Sage.
- Hsieh, M., Hamilton, Z.K. (2014). Predicting success in residential substance abuse interventions: New Jersey's pre release incarceration alternatives. *Criminal Justice Policy Review*, 1-21.
- Imbens, G.W., y Angrist, J.D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2), 467-475.
- Lane, F.C., Gibbs, S. (2015). Propensity score analysis: a secondary data analysis of work-life Policy and performance outcomes. *Advances in Developing Human Resources*, 17(1), 102-116.
- Randolph, J.J., Falbe, K., Kureethara-Manuel, A., & Balloun, J.L. (2014) A step by step guide to Propensity Score Matching. *Practical Assessment, Research and Evaluation*, 19 (18), 1-6.
- Rosenbaum, Rubin (1984) Reducing bias in observational Studies using sub-classification on the propensity score. *Journal of the American Statistical Association*, 79(387), 516-524
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Tanaka, S., Kawakami, K. (2015). Methodological issues in observational studies and non-randomized controlled trials in oncology in the era of big data. *Japanese Journal of Clinical Oncology*, 45(4), 323-327.
- Ten Brummelhuis, L.L., Van Der Lippe, T. (2010). Effective work-life balance support for various household structures. *Human Resource Management*, 49, 173-193.
- Vinha, K. (2006). *A primer on Propensity Score Matching estimators*. Documento CEDE 2006-13. ISSN 1657-7191. Obtenido de: https://economia.uniandes.edu.co/components/com_booklibrary/ebooks/d2006-13.pdf
- What Works Clearinghouse (2012). *Procedures and Standards Handbook (versión 2.0)*. US Department of Education. Institute of Education Sciences.

