Academic achievement and formal thought in engineering students

Stella Maris Vázquez

Hilda Difabio de Anglat

Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires

Argentina

Stella Maris Vázquez. Lacroze 2100. Buenos Aires (1426). Argentina. E-mail: stellavazquez@gmail.com
Hilda Difabio de Anglat. Primitivo de la Reta 522, 2º K, Mendoza (5500). Argentina. E-mail: ganglat@gmail.com

Abstract

Introduction. Research on university-level academic performance has significantly linked failure and dropping out to formal reasoning deficiency. We have not found any papers on formal thought in Argentine university students, in spite of the obvious shortcomings observed in the classrooms. Thus, the main objective of this paper was exploring the relation between formal operational thought level and academic performance in the 1st year of the engineering course.

Method. The study was carried out on a sample of 709 first year engineering students from two cohorts of students attending one of four engineering courses during 2006 and 2007 at an Argentine university. Assessment was taken using Tobin and Capie's (1981) Test of Logical Thinking (TOLT). Data about demographic characteristics and achievement test scores was also collected.

Results. Significant relations were observed between formal thought level and student marks in entrance exams, in the partial algebra, calculus, physics and chemistry exams and in the final chemistry exam.

Conclusion. This research may be useful as a guide for diagnosis and intervention design, since they point to probabilistic and correlational reasoning as the most difficult skills to develop. As the results tend to converge with those of research carried out in other countries, it seems also that the TOLT may be useful in transcultural exploration.

Keywords: Logical thinking- academic achievement-engineering education

Logro académico y pensamiento formal en estudiantes de ingeniería

Resumen

Introducción: La investigación acerca del desempeño en la Universidad relaciona el fracaso y la deserción con las deficiencias en el razonamiento formal. No hemos hallado trabajos sobre el tema, referidos a población argentina, a pesar de las limitaciones observadas en este aspecto. De allí que el objetivo principal de este trabajo sea la presentación de las relaciones halladas entre la operatividad formal y el nivel de desempeño académico en 1er. Año de Ingeniería.

Método: El estudio se hizo con una muestra de 709 estudiantes que cursaban alguna de las cuatro especialidades de Ingeniería en una Universidad argentina. Se aplicó el test de pensamiento lógico-formal (TOLT) de Tobin y Capie. También se recogieron datos sobre características demográficas y rendimiento académico.

Resultados: Se observaron relaciones estadísticamente significativas entre el nivel de pensamiento formal y las calificaciones en el examen de admisión, en los exámenes de Álgebra, Análisis Matemático, Física y Química

Conclusión: Los resultados muestran que el razonamiento probabilístico y correlacional son los menos logrados por los estudiantes de 1er. Año. Los mismos fueron utilizados para el diagnóstico y el diseño de intervención en las actividades tutoriales de la Universidad. La convergencia de estos resultados con los de investigaciones en otros países, muestran que el TOLT puede ser un instrumento válido en la investigación transcultural.

Palabras-clave: Pensamiento formal- rendimiento académico -Ingeniería.

Introduction

Research on university-level academic performance has significantly linked failure and dropping out to formal reasoning deficiency. However, this cognitive dimension has not been particularly studied in the last two decades, in which stress has rather been laid on the role of critical thinking, epistemic beliefs and prior knowledge (Aguilar, Navarro, López & Alcalde, 2002).

According to the Piagetian view of intelligence, formal thought may be defined, from the standpoint of equilibrium, as a stage at which the real is included as a possibility in the broader realm of the possible. This enables the individual to reason from propositions and conceive a complete system of virtual transformations. The latter allows them to foresee the transformations of the real and formulate hypotheses, going beyond what they perceive (Villarroel Villamor, 2009). Thus, intelligence can conceive the logically and physically necessary, and therefore formulate hypotheses, define terms and create models without the support of empirical evidence. From a structural perspective, this gives rise to combination through factor dissociation, which constitutes a fundamental instrument of hypothetical and necessary reasoning. This stage is characterised by the emergence of the 16 binary propositional operations and formal operational schemata: proportion, combination, coordination of reference and motion relativity systems, probability, correlation and multiplicative compensation (Piaget & Inhelder, 1955).

To Piaget, children become capable of formal operations at 11 or 12 years old, or even earlier, as a result of education. However, more recent studies show that this stage is often reached later on, and certain adults (Mwamwenda, 1999, 1993) and university students have not achieved it (Cherian & Kibria, 1999; Godino, Batanero & Roa, 2005; Sutherland, 1982, 1999).

A number of papers demonstrate that formal thought is at the root of mathematics and chemistry problem solving (Aguilar et al., 2002; Boujaoude, Salloum & Abd-El-Khalick, 2004; Lewis & Lewis, 2007), being therefore a predictor of academic achievement and conceptual change in these fields (Oliva, 2003). The Piagetian methodology, based on interviews and real instruments, provides reliable information on thinking levels, enabling researchers to grasp the subjects' thinking processes by considering the strategies used to solve different

situations and the reasons given in each case. The time and expert availability requirements of this approach make it, however, difficult to implement in academic contexts, which has led to the collective and time-limited administration of pen and paper tests (Bart & Schleiman, 1988).

Among them is Tobin and Capie's (1981) Test of Logical Thinking (TOLT), which assesses the level of formal operational thought as described by Piaget (Piaget & Inhelder, 1955). Its multiple-choice items include the justification of the selected answers, which reduces the likelihood of chance success. The instrument covers five skills (proportional, probabilistic, correlational and combinatorial reasoning and variable control), each evaluated by two items. For an answer to be deemed correct both the solution and the reason given must be right. Each item includes a graphic representation of the situation/problem.

It has been used in the present study insofar as it assesses the basic schemata that engineering students must possess in order to approach curriculum subjects. Thus, proportional reasoning is required to understand physical phenomena, to tackle equations, velocity, acceleration and density (Ruiz and Lupiáñez, 2009). Variable identification and control ranks among the main logical skills needed to master physics contents such as the relation between time and movement, and to grasp any experimental design. Similarly, probabilistic and correlational reasoning are at the basis of mathematics, physics and chemistry problem-solving tasks (Adigwe, 1993).

Tobin and Capie (1981) point out that the TOLT has been used chiefly to identify likely successful science and mathematics college students. Thus, research (Mwamwenda, 1993; Valanides, 1996) indicates that formal operations scores correlate significantly with performance in physics, chemistry and biology. More recently, Lewis and Lewis (2007) have found that the TOLT, administered to 3000 1st year college students enrolled in a chemistry course, predicted success rather than failure. Thus, formal thought level would serve to identify barriers to success. These studies further reveal that formal thinking is not universally achieved and quite a number of university-attending adults fail to reach this stage. Between 20 and 25% of the first-year college students in Schwebel's (1975) sample had achieved it, whereas 17% were at the concrete operations level and the remaining 63% on a transition stage. Others (Kuhn & Angelev, 1976; Kuhn, Ho & Adams, 1979), however, show that concrete-thinking level adolescents in their 1st year of college move on to the next stage, after

training, faster than pre-teenage children. From this it is concluded that the former's short-comings might be put down to problem format rather than to lack of competence. In line with these findings, Fishbein (1975, 1987) reports that between 36 and 50% of a sample of 187 12 to 13-year-olds managed to solve combination problems using empirical strategies.

Teachers should acknowledge this situation and focus, beyond algorithmic procedures (Gómez, 2009), on strategies to develop formal operational thinking (Sutherland, 1999). They should be capable of analysing student answers, categorising their mistakes and choosing the most appropriate methods (Roa, Batanero, Rodiño & Cañizares, 1997).

We have not found any papers on formal thought in Argentine university students, in spite of the obvious shortcomings observed in the classrooms. Thus, we have set out to explore the subject, within the framework of a wider project on the variables associated with student drop-out from engineering courses.

We aimed at:

- Determining the validity and reliability of the TOLT in our academic environment.

- Exploring the relation between formal operational thought level and academic performance in the 1st year of the engineering course.

Method

Subjects

Our sample is made up of 709 subjects, 95% of them male, with an average age of 20.7 (sd=3.54). It has been obtained from two cohorts of students attending one of four engineering courses at an Argentine national university during 2006 and 2007. As the wider goal of our project was the study of drop-out reasons, participants were not selected at random.

Instruments

1. To measure formal thought, the Test of Logical Thinking -TOLT- (Tobin and Capie, 1981) has been administered. An answer has been considered correct when both the

solution and the reason given are right. The scores for each of the 5 variables as well as the total scores have been modified to fit a 10-point scale.

2. To measure academic performance we have used the marks for the three partial algebra, calculus, physics and chemistry exams as ordinal variables ranging from 1 to 10. Final exams data are hard to come by, since only a small number of students go in for them immediately after the yearly courses, which led us to create a categorical variable to assess final performance. It has two values: (a) has passed the partial exams and is entitled to take the final exam; (b) has taken the three partial exams but failed to obtain the scores required to take the final exam.

Design and Procedure

Our research design is ecological and correlational because the phenomena are observed as they happen in their natural context in order to evaluate the relationships among the variables in a certain moment. The TOLT has been administered at the beginning of the academic year, during April 2006 and April 2007, in regular class time, in one session, to intact samples by *ad hoc* tutors. Participants were instructed to take their time and read each item carefully, taking an average 40 minutes to complete the TOLT.

Data Analysis

Cronbach's alpha and exploratory factor analyses were used to determine the validity and reliability of the instrument. Item difficulty and discrimination were assessed by Pearson's correlation coefficient and the t-test.

To achieve our second goal we categorised total TOLT scores as well as student entitlement to take the final exams of the four core subjects taught in the first year, as determined by course performance. The relationship between these variables was explored by means of the x^2 and the Kruskall Wallis test. Multiple regression was carried out to establish the extent to which entrance exam marks and total TOLT scores (as a continuous variable) predicted student performance in those subjects.

All statistical analysis was performed through SPSS 11.5.

Results

1. TOLT validity and reliability

Using a 709-subject sample, the test's authors obtained a reliability coefficient of 0.85 for the total score. Subscale internal consistency values were: 0.82 for proportional reasoning and controlling variables, 0.61 for probabilistic reasoning, 0.71 for combinatorial reasoning and 0.56 for correlational reasoning.

Factor analysis of the five subscale scores yielded one factor that accounted for 38% of the variance. Analysis of the ten items included in the subscales showed the five hypothesised factors, explaining 43% of the variance, with loadings between 0.60 and 0.72.

We used the covariance matrix to determine test reliability. Cronbach's alpha was 0.74 for the whole scale, 0.64 for proportional reasoning, 0.80 for controlling variables, 0.70 for probabilistic reasoning, 0.60 for combinatorial reasoning, and 0.51 for correlational reasoning. The last value indicates failure to meet minimum reliability standards.

We used factor analysis to test construct validity. First we applied it to the ten TOLT problems, obtaining a Bartlett's test of sphericity significance of $p < .001$ ($X^2 = 1413.04$, df = 45). Thus the null hypothesis was rejected that the sampling matrix had been developed from a population in which the variables did not correlate. The KMO value of 0.72 suggested that the variables could be explained by fewer factors. Our analysis yielded five factors accounting for 75% of the variance

Table 1. Pattern matrix (a)

Components	1	2	3	4	5
Proportional Reasoning 1	0.87	-0.15	0.06	-0.18	-0.24
Proportional Reasoning 2	0.85	-0.21	0.23	-0.28	-0.22
Variable control 2	0.16	-0.91	0.14	-0.25	-0.25
Variable control 1	0.20	-0.90	0.15	-0.28	-0.30
Combinatorial Reasoning 1	0.14	-0.08	0.85	-0.13	-0.16
Combinatorial Reasoning 2	0.12	-0.20	0.81	-0.22	-0.27
Probabilistic Reasoning 1	0.21	-0.23	0.19	-0.88	-0.28
Probabilistic Reasoning 2	0.25	-0.29	0.17	-0.87	-0.33
Correlational Reasoning 2	0.183	-0.201	0.212	-0.259	-0.86
Correlational Reasoning 1	0.269	-0.321	0.223	-0.338	-0.76

Extraction method: Principal Components Analysis. Rotation method: Oblimin with Kaiser Normalisation.
(a) Convergence is reached at 6 iterations.

As the factors were correlated, an Oblimin rotation was used. The table shows that factor loadings ranged from 0.81 to 0.91, except in the case of the second correlation problem, a loading value of -0.764 being at any rate acceptable.

Given this correlation, we conducted a second order factor analysis on the five subscales, obtaining one factor that explained 41% of the variance, with a KMO of 0.75 and a significant Bartlett's test of sphericity ($X^2 = 369.70$, df = 10, p<0.001).

Table 2. Components Matrix (a)

Component	1
Correlational Reasoning	0.73
Probabilistic Reasoning	0.69
Variable control	0.63
Proportional Reasoning	0.60
Combinatorial Reasoning	0.53

Extraction method: Principal Components Analysis.
(a) 1: extracted components

These results are similar to the ones reported by the TOLT authors, our first analysis accounting for a greater percentage of variance. Therefore, we were able to conclude that the test has good internal consistency and construct validity and might be administered to our engineering student population to assess formal operational thought. The coefficients of correlation between subscales are significant but moderate, with significant differences between

subscale totals. This indicates that the skills evaluated have both common and distinctive characteristics, which could be gathered also from the order of difficulty, as will be seen next.

2. Item discriminating power and difficulty level

A discriminating index has been used to assess item quality, calculated through:

(1) Contrast between the highest and lowest 27%, a percentage that enables the creation of extreme groups that are as large and different from each other as possible (Vianna, 1983). Student's t-test has been used to establish the significance of the difference between group means for each item.

(2) The biserial correlation coefficient, which provides an estimate of Pearson's product moment correlation between total test scores and the hypothetical continuum of the items as dichotomised into correct and incorrect responses (Backhoff, Larrazolo & Rosas, 2000). It determines whether a certain item has been correctly answered by the 'right' people, eloquently revealing the predictive validity of the test (Henrysson, 1971, quoted by Backhoff, Larrazolo & Rosas). This method has the advantage of considering the whole sample, as opposed to the 54% taken into account by the former one.

The statistical significance of both kinds of index for a given item has constituted our criterion for psychometric quality. We obtained the following results:

Table 3. Item discrimination

Item	Student's t	Biserial correl.
1.	9.52	0.44**
2.	11.29	0.54**
3.	19.32	0.60**
4.	16	0.57**
5.	17.9	0.58**
6.	18.9	0.61**
7.	20.4	0.60**
8.	13.4	0.54**
9.	14.4	0.44**
10.	12.7	0.52**

** All the coefficients are significant
at 0.001 level (two-tailed).

Student's t values for the 10 items are statistically significant (p< 0.001). These points to a very significant difference –not due to chance- between the mean scores of the top and bottom 27% of the sample for each item. The biserial correlation coefficient also indicates significant statistical relations for all items.

Item difficulty is in inverse proportion to the ratio of correct answers to the total number of responses: the greater this ratio, the less difficult the item. To calculate it, the number of students who answered it rightly has been divided by the number of students in the sample.

Thus, the average difficulty level of the TOLT is 0.57, which meets the requirements of statistical theory (between 0.5 and 0.6). It has been further recommended that instruments for educational purposes should consist mainly of items within the 0.20 to 0.80 difficulty interval, with only a limited number of items outside it (Vianna, 1983). Backhoff, Larrazolo and Rosas (2000) suggest that very easy (DL 0.87 to 1) items should amount to 5% of the test, easy (DL 0.74 to 0.86) ones to 20%, moderately difficult (DL 0.53 to 0.73) ones to 50%, difficult (DL 0.33 to 0.52) ones to 20% and very difficult (DL 0 to 0.32) items to 5%. According to our results none of the TOLT items is very easy, only one is easy (n 1), 50% are moderately difficult (n 2, 3, 4, 8, and 9), 30% are difficult (n 5, 7, and 10), and one (n 6) is very difficult (Figure 1).

Table 4. Item difficulty

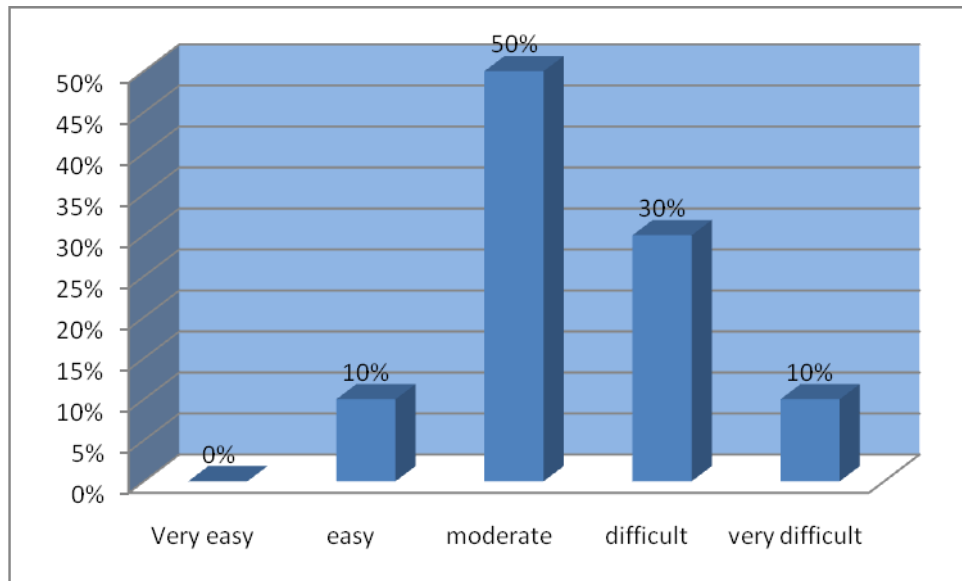| Item	Difficulty Level
1.	0.84
2.	0.67
3.	0.59
4.	0.64
5.	0.39
6.	0.32
7.	0.44
8.	0.53
9.	0.63
10.	0.53

Figure 1. Item frequency for difficulty level

3. Formal thought level

The subjects' scores have been modified to fit a 10-point scale, with the following results:

Table 5. Descriptive statistics

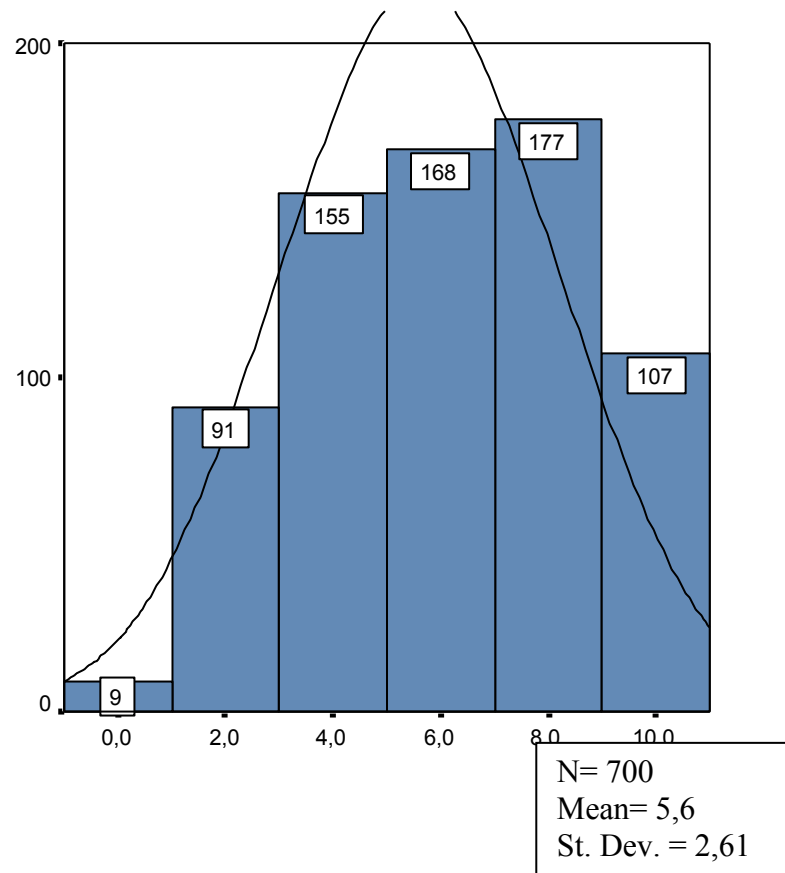| N	709
Mean	5.60
Skewness	-0.108
Skewness error	0.092
Kurtosis	-0.926
Kurtosis error	0.183

Figure 2. TOLT score distribution

The distribution is asymmetric within normal parameters. The negative kurtosis can be explained by the make-up of the sample, since we administered to first-year university students a test originally intended for 12- to 18-year-olds, i.e. for secondary school pupils. Similar results have been reported with US students in their first year of college (Lewis & Lewis, 2007), although the mean was 6.8.

In considering subscale difficulty, we have noticed that the subskills with the highest means are proportional reasoning ($\overline{X}$ = 7.55) and variable control ($\overline{X}$ = 6.15), whereas proabilistic ($\overline{X}$ = 3.57) and correlational ($\overline{X}$ = 4.95) reasoning have the lowest. This partially corresponds with Oliva's (1999) findings in a sample of 15- and 16-year-old Spanish students, for which the mean score was 2.87. Valanides (1997) administered the test to 469 Cypriot students in their last year of secondary school; 50 % of those enrolled in the classical studies, business studies and foreign languages sections were at the concrete thought stage, and only 2% had, strictly speaking, developed formal thinking. For those enrolled in the science sec-

tion, the percentages were, respectively, 5 and 30%. Only 17% of the whole sample was at the formal operations stage. Proportional reasoning and variable control problems had the highest scores, while probabilistic reasoning and the second combinations problem had the lowest. The mean for male subjects was 4.80. We can therefore safely assume that the level of difficulty is constant.

In line with the criterion applied by other researchers using the TOLT (Aguilar et al., 2002; Raviolo, Siracusa, Herbel & Schnersch, 2000; Yenilmez, Sungur & Tekkaya, 2006), we have categorised the scores, generating a trichotomous variable. Score distribution has been taken into account.

Table 6. Categorised scores

Score	Frequency	%	Accum. %
:≤4	256	36.1	36.1
2:4-6	168	23.7	59.8
3:≥ 7	285	40.2	100
Total	709	100	

The first category corresponds to the concrete operations level, the second to the transition stage and the third to formal thought (Frear & Hirschbuhl, 1999). Thus, 33% of our sample is at the concrete thinking stage, and only 39% have developed formal operational schemata.

4. Formal thought and academic achievement

To evaluate academic performance we have focused on the most difficult subjects taught in the first year of the engineering course -algebra, calculus, physics and chemistry-, using students' marks in the partial and final exams. A dichotomous categorical variable measures final performance through student entitlement to take the final exams. A chi-square test on these data indicates a significant relation between formal thought level and final performance in calculus ($X^2 = 15.86$, df = 2, p< .02), physics ($X^2 = 7.89$, df = 2, p< .02), chemistry ($X^2 = 9.25$, df = 2, p< 01) and algebra ($X^2 = 22.08$, df = 2, p< .01). The TOLT has proved more useful in identifying high achievers: 66% of the students that passed the partial exams have reached the formal operations stage. In all the subjects, at the third TOLT level we find the highest percentage of individuals that passed the partial exams and

the lowest percentage of those that did not. The opposite happens in algebra and calculus at the first level, which coincides with the concrete thought stage, as can be seen in the following tables.

Table 7. Calculus results

CALCULUS	TOLT CATEGORY		
	<=4	4-6	≥ 7
Failure	57.0%	49%	34%
Success	43.0%	51%	66%
Total	100.0%	100.0%	100.0%

Table 8. Chemistry results

CHEMISTRY	TOLT CATEGORY		
	<=4	4-6	≥ 7
Failure	41%	45%	34%
Success	59%	55%	66%
Total	100.0%	100.0%	100.0%

Table 9. Physics results

PHYSICS	TOLT CATEGORY		
	<=4	4-6	≥ 7
Failure	49%	45%	32%
Success	51%	55%	68%
Total	100.0%	100.0%	100.0%

Table 10. Algebra results

ALGEBRA	TOLT CATEGORY		
	1.00	2.00	3.00
Failure	67%	55%	36%
Success	33%	45.%	64%
Total	100.0%	100.0%	100.0%

As three of the variables were not homoscedastic, we used the Kruskall Wallis test to analyse the relation between formal operations level and student performance in the partial and final exams already considered, as well as in the entrance (physics and mathematics) exams.

Table 11. Kruskall Wallis Test

	Math en	Physics en	Alge par	Alge final	Calcul par	Calcul final	Physics par	Phys final	Chemist par	Chem final
X^2	51.96	36.35	14.12	2.10	13.22	4.04	18.75	3.92	28.18	14.02
D Df	2	2	2	2	2	2	2	2	2	2
sign.	0.00	0.00	0.00	0.35	0.00	0.13	0.00	0.14	0.00	0.00

The results partially confirm our hypotheses. Significant relations can be observed between formal thought level and student marks in both entrance exams, in the partial algebra, calculus, physics and chemistry exams and in the final chemistry exam. In the remaining cases, the lack of such relation points to the problem that has prompted our study and manifests itself in the small number of students that go in for the final exams. Between 45% and 55% of the students enrolled in each course pass the partial exams, and 50% of these take the final one. Only between 16 and 21% of those originally enrolled pass it. Poor performance in these disciplines becomes an obstacle to the normal pursuit of the engineering course. It leads to high drop-out rates and increases the number of first-year students because of the need to repeat the basic courses.

It is worth noticing, however, that between 60 and 66% of the students that pass the final basic subjects exams have reached the formal operations level, an equal proportion of the remaining 40% being at the two previous stages. These data would bear out our hypothesis.

Performance differences occur between TOLT levels 1 and 3, except in the case of chemistry, for which there are significant differences between the three formal operations levels.

Linear regression was used to ascertain whether entrance exam and TOLT scores predict later achievement. It was established that both factors account for 47% of the variance in final chemistry performance (R^2 0.47, F = 11.92, p < 0.001), and 26% of the variance in physics performance. Performance in the remaining subjects behaves independently. These findings will bear on our eventual approach to the problem we have set out to solve.

Conclusions

The TOLT has proved a valuable and reliable instrument for exploring the level of formal operational thought reached by first-year engineering students. It has enabled us to place our subjects at one of three levels and establish significant relations between them and academic performance.

The mean for our sample is, nevertheless, lower than the one reported in similar studies. Moreover, a certain natural selection would be expected in the case of engineering students, which underscores the difference. The fact that 36 % of them have been found to be at the concrete thinking stage suggests the need for remedial intervention during the entrance course. The results of our research may be useful as a guide for diagnosis and intervention design, since they point to probabilistic and correlational reasoning as the most difficult skills to develop. However, both the variance in performance explained by formal operational thought and the entrance examination performance would lead us in other directions. Other variables might be at stake, such as type of motivation, study strategies and learning orientation. We should also think of the instructional methods used in first-year university courses[1]. Other authors (Roa, Batanero, Rodino & Cañizares, 1997) have shown that even advanced mathematics students experience difficulty in solving combinatorial problems. Case analysis has revealed that low achievers make little use of tree diagrams or strategies such as fixing variables or dividingup problems as steps towards generalisation (Eizenberg & Zaslavsky, 2004).Instead of suggesting these strategies, teachers often focus on algorithms and operation definitions (Boujaoude et al. 2004), failing to encourage resourceful thinking. The schemata underlying students' intuitive solutions are often neglected, when they should be discovered and taken into consideration until the appropriate schemata become intuitive (Fischbein, 1987; Fischbein & Grossman, 1997). These and other analogous studies (Sutherland, 1999) might provide useful guidance for teachers' work.

Remedial intervention does not amount to turning university into one more stage of secondary education. As has been demonstrated (Kuhn et al. 1979), university students at the concrete thought level pass on to the next stage faster than those in their last-year of secon-

---

[1] We have explored these variables elsewhere. A discussion of the results would go beyond the limits of this paper.

dary schooling. This points to the existence of a transitional stage that university teachers would be wise not to ignore.

We hope that our work might contribute to the diagnosis of the specific learning difficulties experienced by the population represented in our sample, as well as to the ensuing instructional intervention. This could especially be the case in Argentina, where we have not found any other papers dealing with this issue. Since our results and those of research carried out in other countries with the same instrument tend to converge, we also think our project may play a part in its trans-cultural exploration.

References

Adigwe, J. C. (1993). Some correlates of Nigerian students' performance in chemical problem-solving. *Research in Science & Technological Education*, *11*(1), 39-48.

Aguilar Villagrán, M., Navarro Guzmán, J. I., López Pavón, J. M., & Alcalde Cuevas, C. (2002). Pensamiento formal y resolución de problemas matemáticos. *Psicothema*, *14*(2), 382-286.

Backhoff, E., Larrazolo, N., & Rosas, M. (2000). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación Educativa, 2*(1), 11-29.

Bart, W., & Schleisman, K. (1988). Testing formal reasoning. *Applied Measurement in Education*, *1*(2), 189-203.

Boujaoude, S., Salloum, S., & Abd-El-Khalick, F. (2004). Relationships between selective cognitive variables and students' ability to solve chemistry problems. *International Journal of Science Education*, *26*(1), 63–84.

Cherian, V., & Kibria, G. (1999). Formal operational reasoning in African university students. *The Journal of Psychology*, *122*(5), 487-498.

Eizenberg, M., & Zaslavsky, O. (2004). Students´ verification strategies for combinatorial problems. *Mathematical Thinking and learning*, *6*(1), 15-3.

Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. Dordretch: Reidel.

Fischbein, E. (1987). *Intuition in science and mathematics*. Dordretch: Reidel.

Fischbein, E., & Grossman, A. (1997). Schemata and intuitions in combinatorial reasoning. *Educational Studies in Mathematics*, *34***,** 27–47.

Frear, V., & Hirschbuhl, J. J. (1999)**.** Does interactive multimedia promote achievement and higher level thinking skills for today's science students? *British Journal of Educational Technology*, *30*(4), 323–329.

Godino, J., Batanero, C., & Roa, R. (2005). An onto-semiotic analysis of combinatorial problems and the solving processes by university students. *Educational Studies in Mathematics*, *60*, 3–36.

Gómez, P. (2009). Procesos de Aprendizaje en la Formación Inicial de Profesores de Matemáticas de Secundaria. *Electronic Journal of Research in Educational Psychology*, *17*, *7*(1), 471-498.

Kuhn, D., & Angelev, J. (1976). An experimental study of the development of formal operational thought. *Child Development*, *47,* 697-706.

Kuhn, D., Ho, V., & Adams, C. (1979). Formal reasoning among pre- and late Adolescents. *Child Development*, *50*, 1128-1135.

Lewis, S., & Lewis, J. E. (2007). Predicting at-risk students in general chemistry: comparing formal thought to a general achievement measure. *Chemistry Education Research and Practice*, *8*(1), 32-51.

Mwamwenda, T. (1993). Formal operations and academic achievement. *The Journal of psychology*, *127*(1)*,* 99-103.

Mwamwenda, T. (1999). Undergraduate and Graduate Students' Combinatorial Reasoning and Formal Operations.*Journal of Genetic Psychology*, *160*(4), 503-507.

Oliva, J. (1999). Structural patterns in students' conceptions in mechanics. *International Journal of Science Education*, *21*(9), 903– 920.

Oliva, J. (2003). The structural coherence of students' conceptions in mechanics and conceptual change. *International Journal of Science Education*, *25*(5), 539-561.

Piaget, J., & Inhelder, B. (1955). *De la Logique de l′enfant á la logique de l′adolescent*. Paris: Presses Universitaires de France.

Raviolo, A., Siracusa, P., Herbel, M., & Schnersch, A. (2000). Desarrollo de razonamientos científicos en la formación inicial de maestros. *Revista Interuniversitaria de Formación de Profesorado*, *38*, 129-140.

Roa, R., Batanero, C., Rodino, J. D., & Cañizares, M. J. (1997). Estrategias en la Resolución de Problemas Combinatorios por Estudiantes con preparación matemática Avanzada**,** *Epsilon*, *36,* 433-446.

Ruiz, E. & Lupiáñez, J. (2009). Detección de obstáculos psicopedagógicos en la enseñanza y el aprendizaje de los tópicos de razón y proporción en alumnos de sexto grado de Educación Primaria. *Electronic Journal of Research in Educational Psychology*, *17*, *7*(1), 397-424.

Schwebel, M. (1975). Formal operations in first-year college students. *The Journal of Psychology*, *91*, 133-141.

Sutherland, P. (1982). An expansion of Peel's describer-explainer stage theory. *Educational Review*, *34*(1), 69-76.

Sutherland, P. (1999). The application of Piagetian and neo-Piagetian ideas to further and higher education. *International Journal of Lifelong Education*, *18*(4), 286-294.

Tobin, K., & Capie, W. (1981). The Development and Validation of a Group Test of Logical Thinking. *Educational and Psychological Measurement*, *41*(2), 413-419.

Valanides, N. (1996). Formal reasoning and science teaching. *School Science and Mathematics*, *96*(2), 99-107.

Valanides, N. (1997). Cognitive abilities among twelfth-grade students: Implication for science teaching. *Educational Research and Evaluation*, *3*(2), 160-186.

Vianna, H. (1983). *Los tests en educación*. Pamplona: EUNSA.

Villarroel Villamor, J. (2009). Origen y desarrollo del pensamiento numérico: una perspectiva multidisciplinar. *Electronic Journal of Research in Educational Psychology*, *17*, *7*(1), 555-604.

Yenilmez, A., Sungur, S., & Tekkaya, C. (2006). Students' achievement in relation to reasoning ability, prior knowledge and gender. *Research in Science & Technological Education*, *24*(1), 129–138.