Published in final edited form as:

J Chem Educ. 2010 March ; 87(5): 541–546. doi:10.1021/ed8001754.

LOGICAL REASONING ABILITY AND STUDENT PERFORMANCE IN GENERAL CHEMISTRY

Dr. Lillian Bird

Department of Chemistry, University of Puerto Rico at Río Piedras, PO Box 23346, San Juan, PR 00931-3346

Abstract

Logical reasoning skills of students enrolled in General Chemistry at the University of Puerto Rico in Río Piedras were measured using the Group Assessment of Logical Thinking (GALT) test. The results were used to determine the students' cognitive level (concrete, transitional, formal) as well as their level of performance by logical reasoning mode (mass/volume conservation, proportional reasoning, correlational reasoning, experimental variable control, probabilistic reasoning and combinatorial reasoning). This information was used to identify particular deficiencies and gender effects, and to determine which logical reasoning modes were the best predictors of student performance in the general chemistry course. Statistical tests to analyze the relation between (a) operational level and final grade in both semesters of the course; (b) GALT test results and performance in the ACS General Chemistry Examination; and (c) operational level and student approach (algorithmic or conceptual) towards a test question that may be answered correctly using either strategy, were also performed.

Keywords

First Year Undergraduate/General; Chemical Education Research; Problem Solving; Testing/Assessment; Gases

Introduction

Since the 1970's, the identification of predictors of academic performance in science courses has been an object of study among science education researchers (1-20). During past years, student performance in science courses has been analyzed on the basis of high school experiences (15), precalculus grades (8), math diagnostics (10), verbal and math components of the SAT (5,15) and ACT (8), mental capacity measured as M-Demand through the Figural Intersection Test (19), formal operational reasoning measured by using the Test of Logical Thinking (19) and the Group Assessment of Logical Thinking (GALT) test (5,6,10), and disembedding ability measured through the Group Embedded Figures Test (19), among many others. Researchers have used these parameters, individually as well as combined, to try to find a true predictor of achievement in undergraduate science courses.

Advances in the field of cognitive psychology have emphasized the importance of cognitive skills as key elements for the acquisition of knowledge in introductory science courses (21). These advances have prompted researchers to give more serious thought to cognitive factors for predicting student achievement at the undergraduate level. To better understand our own science student population and contribute to this discussion, on March 2007 a Spanish

version of the GALT test was administered to students enrolled in the General Chemistry course at the University of Puerto Rico in Río Piedras (UPR-RP). This test is a 12-item instrument developed by Roadranga *et al* (22) to measure logical reasoning skills in pre-college and college level students. The test includes questions related to mass and volume conservation, proportional reasoning, correlational reasoning, control of experimental variables, probabilistic reasoning and combinatorial reasoning. The first of these skills (mass/volume conservation) is typically mastered at the concrete operational level, whereas all others correspond to the formal domain. Results of this 12-item test are used to determine the operational level of the responder. GALT results between 0-4 are characteristic of concrete thinkers, 5-7 of individuals in a transitional stage, and 8-12 of formal thinkers (16,22). Studies conducted at institutions where the English version of this test was administered show that, in terms of operational level, most introductory college students seem to be at the transitional or formal operational stages (4,5,16,23).

An analysis of our students' ability to operate at a given level and its relation to academic performance in the General Chemistry course will be presented in this paper. Performance in the GALT test and in individual reasoning skills by gender and by operational level, as well as findings with regards to the relation between operational level and student performance in the ACS General Chemistry Examination and between operational level and student approach (algorithmic or conceptual) towards a partial exam question that may be answered correctly using either strategy, will also be discussed.

Methodology

The population studied was comprised of 466 students enrolled in General Chemistry (CHEM 3001-3002) at UPR-RP who took the GALT test and completed both semesters of the course during academic year 2006-2007. Of these, 66.3% were female; 48.9% were freshmen; and 26.6% took the course in the *Personalized System of Instruction* (PSI) format whereas the rest attended traditional lecture sessions. Most students (81.3%) belonged to the Faculty of Natural Sciences, 17.4% to the College of Education and the remaining 6 students were registered in other colleges. More than 99% of the students were Hispanic, almost all Puerto Rican.

The 12-item GALT test was translated into Spanish (with permission from its authors) and administered to students during the first laboratory session of the Spring semester of academic year 2006-2007. Students were allowed 45 minutes to complete the 12-item test. Ten of these items are comprised of two multiple choice questions, the first to select the proper answer to a particular situation and the other to select the rationale behind this answer. To determine internal reliability and compare student performance in different logical reasoning skills, individual results for each test item were recorded. All data was saved in Excel (24) and exported to SPSS 15.0 (25) to perform statistical analyses.

All student data was obtained upon approval by the Institutional Committee for the Protection of Human Subjects in Research. Student responses to a partial exam question were analyzed by the author by examining the tests of students enrolled in several sections of the course.

Results and discussion

Internal reliability of the GALT test

Internal reliability coefficients convey the degree of coherence among items in an instrument intended to measure a given construct. For psychometric instruments, internal reliability is most commonly determined by calculating Cronbach's alpha (26,27). The

magnitude of Cronbach's alpha, however, is not only dependent on the degree of correlation among items but also varies with test length, with longer tests yielding higher values than shorter ones. Thus for short tests, the Spearman-Brown prophecy, which allows the researcher to predict the internal reliability that would result from an increase in test length, is frequently calculated (26). In practical terms, reliability standards for psychometric instruments were set by Nunnally in his seminal work on *Psychometric Theory*. According to Nunnally, a reliability value of 0.70 is sufficient for preliminary work on predictor tests, whereas for basic research on psychometric instruments reliability values need not exceed 0.80 (26).

For the Spanish version of the GALT test administered to students at UPR-Río Piedras, Cronbach's alpha yielded a value of 0.69, and the Spearman-Brown prophecy calculated for a "21-item" test was 0.79. These values are in the same range as those obtained by Bunce *et al* (5), which were 0.62 and 0.74, respectively, using the English version of the test with college level students. They are lower, however, than the alpha value of 0.85 reported by Roadrangka *et al* (28), as well as by Bitner (29), all of whom worked with pre-college students.

In order to compare the internal reliability of the test administered at UPR-RP to that obtained by Niaz and Robinson (4), the Guttman split-half reliability coefficient was also determined. This reliability coefficient, which indicates the degree of correlation between the two halves of a test, is an alternate measure of internal consistency (26). It is important to point out that, for the GALT test, the particular split-half item division is important because the first two items address mass/volume conservation, the next two deal with proportions, items 5 and 6 are based on control of experimental variables, and so on, making the first half of the test different from the second half, yet the odd half equivalent to the even half. Taking this into account, a Guttman reliability coefficient of 0.74 was obtained by comparing odd-numbered items with their even-numbered counterparts. This value is higher than that obtained by Niaz *et al* (4) for the same test, which was 0.63.

Student distribution by operational level

Based on GALT test results, the operational level of students enrolled in General Chemistry at UPR-RP during academic year 2006-07 was determined. Students with scores ranging from 0-4 were considered to be in the concrete operational level, from 5-7 in the transitional stage, and from 8-12 in the formal operational level (16,22). On the basis of these ranges, 19% of the students were in a concrete level, 40% were in a transitional stage, and 41% had reached the formal operational level. Because the GALT test was administered at the beginning of the second semester, these percentages do not include students who failed or withdrew from the course during the first semester of the course. It would be reasonable to speculate that this excluded cohort could further increase the percent of concrete and transitional operators, at the expense of that of formal thinkers. In any case, the fact that at least 59% of all students enrolled in this course fall below the formal operational level is quite troubling, because mastery of most topics covered in this and subsequent chemistry courses requires formal reasoning skills.

The above findings are comparable to those obtained by McConnell *et al* (16) for students in introductory geoscience courses, which were found to be 24% concrete, 33% transitional and 43% abstract thinkers on the basis of the same test and score ranges. They differ, however, from those obtained by McKinnon and Renner (1), who describe their college students as 50% concrete, 25% post-concrete and 25% formal using another instrument.

Gender and logical reasoning ability

A t-test for independent samples was performed to compare GALT test scores among students of different gender. A significant effect for gender was observed, $t(464) = 5.09$, $p < .001$ (two-tailed), with male students obtaining higher scores. Similar findings regarding gender differences in test scores, not necessarily using GALT, have also been reported by McKinnon and Renner (1), and Shibley *et al* (14). A significant effect for gender was accordingly observed with respect to operational level, $\chi^2(2, N = 466) = 20.16$, $p < .001$, with male students being more likely to be at a formal stage.

GALT scores and other parameters

No significant difference ($p > .05$) was found among students in the traditional lecture vs. PSI sections of the course. Freshmen students, however, did better than upperclassmen, $t(464) = 6.15$, $p < .001$ (two-tailed). This seemingly paradoxical finding as, according to Piaget's levels of development, logical reasoning ability should increase with age (12), may be explained by the fact that only those students with high GPA and College Board scores were allowed to take the general chemistry course during their freshman year, thus raising the stakes for this cohort. A significant difference in GALT test scores, $t(458) = 6.71$, $p < .001$, was also found between students belonging to the Faculty of Natural Sciences when compared to those enrolled in the Faculty of Education. This finding may be explained on the basis of differences in admission criteria for both faculties, the Faculty of Natural Sciences having more rigorous entrance requirements.

Student performance by logical reasoning mode

Student scores in each of the six logical reasoning modes were analyzed in order to determine whether students perform at the same level in each of these modes or if differential performance is observed. Results in Table 1 indicate that, of a maximum score of 2.00, students performed significantly better in items related to combinatorial reasoning, mass/volume conservation, and control of experimental variables; than in those related to probability, proportional reasoning, and correlation.

These results may be compared to those obtained by Bitner (29) using the same test. Bitner found that students of grades 9 through 12 performed significantly better in items related to mass/volume conservation ($M = 1.42$, $SD = 0.68$), control of experimental variables ($M = 0.91$, $SD = 0.76$), and combinatorial reasoning ($M = 0.75$, $SD = 0.75$); than in those related to proportional reasoning ($M = 0.63$, $SD = 0.73$), probabilistic reasoning ($M = 0.53$, $SD = 0.86$), and correlational reasoning ($M = 0.19$, $SD = 0.42$). Although not in the same order as our findings, her results corroborate the students' difficulty with respect to correlational reasoning and, to a lesser extent, with probabilistic and proportional reasoning.

Student performance in logical reasoning modes by gender

Results of the t-tests performed to compare mean scores in each of the modes by gender are summarized in Table 2. Worth mentioning are the findings that male students do markedly better in proportional and probabilistic reasoning than female students ($M_{\text{male}} = 1.15$: $M_{\text{female}} = 0.74$, and $M_{\text{male}} = 1.26$: $M_{\text{female}} = 0.83$, respectively); and that both male and female students do poorly when asked to identify a possible correlation between two variables ($M_{\text{male}} = 0.82$: $M_{\text{female}} = 0.66$). No statistically significant gender difference ($p > .05$) was found with respect to experimental variable control.

Student performance in logical reasoning modes by operational level

Results of the ANOVA test used to compare means for each logical reasoning mode among students of different operational level are summarized in Table 3. A *post-hoc* comparison

test using the Bonferroni correction confirmed that differences in mean score by operational level were significant at the $p < .05$ level.

Based on differences in relative mean score, these results suggest that proportional, probabilistic and correlational reasoning skills are the domain of the formal operational stage; whereas skills in mass/volume conservation, control of experimental variables and combinatorial reasoning may be mastered by students operating at the transitional or concrete levels. As mentioned above, it is evident that correlational reasoning ability is the hardest to attain, even for students operating at the formal level.

Operational level and performance in the General Chemistry course

Final grade in General Chemistry for students of different operational level was analyzed. Results indicate that this grade differed significantly by operational level; both for the first semester of the course, $\chi^2(8, N = 466) = 52.89, p < .001$, as well as for the second semester, $\chi^2(8, N = 466) = 52.48, p < .001$. Figure 1 shows the grade distribution by operational level for students who passed General Chemistry II. A similar graph was obtained for the first semester of the course. From inspection, it is evident that, in terms of final grade, the mode for students operating at a formal level is A, for those at a transitional level is B and for students at a concrete level is C.

GALT results and student performance in the General Chemistry course

A moderate correlation, a Pearson's R of 0.338 ($df = 464, p < .001$, two-tailed) for General Chemistry I and of 0.318 ($df = 464, p < .001$, two-tailed) for General Chemistry II, was found between GALT test results and the students' final grade in both semesters of the course.

Logical reasoning modes and student performance in the General Chemistry course

As mentioned above, most students in the sample show less proficiency in correlational, proportional and probabilistic reasoning than in other logical reasoning modes. This, however, does not necessarily imply that these three modes are the best predictors of performance in the course or that other modes are less important. In order to find out to what extent each of the six logical reasoning modes is able to predict student achievement in the course, a multiple regression was performed. The results obtained indicate that, of the six reasoning modes, probabilistic reasoning is the best independent predictor of student performance in both semesters of the course. Table 4 summarizes the correlation data for all modes.

Further studies are being conducted to compare these results with those for other introductory science courses taken by the same student population. However, in view of the fact that Bitner's (29) as well as this author's findings suggest that some logical reasoning skills are acquired after others have been mastered, this finding should not be used to discard any mode on the basis of its significance for success in the course.

Operational level and problem solving ability

Algorithmic vs. conceptual problem solving ability of general chemistry students has been the subject of many studies in past years (30-35). Nurrenbern and Pickering (30) compared student responses to algorithmic and conceptual questions, and found that many students who correctly answered algorithmic problems did not understand the chemical concepts behind these problems. Lythcott (31) suggested that differences in cognitive development, among other factors, may explain these findings. Based on the work of Tobias (32) and using questions similar to those of Nurrenbern and Pickering (30), Nakhleh (33,34) categorized students as high (H) or low (L) algorithmic (A) and conceptual (C) problem

solvers, and encountered very few LA-HC students (only 5%). Through interviews she also discovered that most students who had correctly responded to the conceptual questions had solved them by applying algorithmic strategies (34). Zoller *et al* (35) found that, whereas most students showed proficiency in the use of algorithms and low order cognitive skills, a significantly lower number could answer conceptual questions. More recently, using Nurrenbern and Pickering's questions, Cracolice *et al* (36) correlated reasoning ability (top 1/3 reasoners vs. bottom 1/3 reasoners on the basis of the Classroom Test of Scientific Reasoning) with success in algorithmic vs. conceptual questions, and concluded that "a significant fraction of our students have no choice other than to be algorithmic problem solvers because their reasoning skills are not sufficiently developed to allow them to successfully solve conceptual problems."

In order to determine whether there is a relation between operational level and problem solving ability, the students' approach towards a partial exam question that could be answered correctly either by algorithmic problem solving or by conceptual reasoning was analyzed by inspecting their responses to the following problem:

Individual 0.200 g samples of each of the following gases were placed in four separate 1.00 L stoppered flasks at 298 K. In which flask do you expect the gas to exert more pressure? Explain your answer

Flask:	1	2	3	4
Gas sample in flask:	CH ₄	Ne	N ₂	CO ₂
Molar mass (g/mol):	16.0	20.2	28.0	44.0

As is evident, all data required to calculate the pressure exerted by each gas is provided. However, given the fact that the temperature, volume and mass are exactly the same for the four gases, no calculations are necessary. The flask containing more moles (i.e., that containing the gas of lower molar mass) will exert more pressure.

Of a total of 106 students whose responses were analyzed and who answered the question correctly, 31% calculated the pressure for each gas and justified their answer in terms of the results obtained (algorithmic approach); 42% did all the calculations, analyzed the results and proceeded to give the correct answer in terms of logical reasoning (more moles of gas, thus higher pressure), which represents the mixed algorithmic/conceptual approach; and 26% gave the correct answer exclusively in terms of logical reasoning (more moles of gas, thus higher pressure) without doing any calculations (conceptual approach). It should be noted that the course instructors did not expect students to use the algorithmic approach to answer this question, to the extent that very little space was provided for students to write their answer. In the case of algorithmic/conceptual answers, the lack of space made it easier to determine what had come first: the algorithmic calculations or the logical reasoning component of the answer.

The results obtained, which are summarized in Table 5, indicate that student approach towards the question differed significantly by operational level, with students at a formal stage being more likely to choose a conceptual approach than students at the transitional or concrete operational levels. Values obtained for the chi-square test were as follows: $\chi^2(4, N = 106) = 11.90, p < .05$.

It should be noted that 33 out of 78 students (42%) who initially answered the question in an algorithmic fashion never provided the expected conceptual answer. All these findings

support the conclusion of Cracolice *et al* (36) with regards to the poor development of reasoning skills as a determining factor for the students' inability to solve conceptual problems in chemistry; and are in agreement with Zoller *et al*'s conclusion that "success on algorithmic questions on exams does not imply success on conceptual questions" (35).

Operational level and performance in the ACS General Chemistry Examination

In order to determine a potential relation between a student's logical reasoning ability and his/her performance in a national chemistry exam, the mean scores for students of different operational level in the ACS General Chemistry Examination were compared. The results of the ANOVA test are summarized in Table 6. A *post-hoc* multiple comparisons test using the Bonferroni adjustment confirmed that differences in mean score were significant at the $p < .01$ level.

These results indicate that there is a statistically significant difference in mean score among students of different operational level, and once again suggest that the attainment of logical reasoning skills is an essential element for mastery of general chemistry concepts and problem solving skills.

Conclusions

On the basis of the aforementioned results we may conclude that, for the student population assessed: (1) most students enrolled in general chemistry (59%) have not reached the formal operational stage; (2) logical reasoning ability (as measured using the GALT test) is a valid predictor of student performance in both semesters of the course; (3) of the six logical reasoning modes, most students show marked deficiencies in proportional, probabilistic and correlational reasoning; (4) among the six logical reasoning modes, probabilistic reasoning is the single best predictor of student performance in general chemistry; (5) student approach towards a chemistry exam question that may be answered correctly using either an algorithmic or a conceptual pathway varies significantly with the student's operational level, with formal thinkers having a stronger tendency to apply the conceptual approach; and (6) students at a formal operational stage perform significantly better in the ACS General Chemistry Examination than students operating at lower levels.

All of these findings imply that logical reasoning skills are essential for student mastery of many of the concepts and more complex problem solving strategies required to succeed in general chemistry. The fact that most students taking this course have not reached the formal stage means that, although they may master the algorithmic component and some basic concepts of the course, many will not be able to interpret their results on the basis of chemical behavior, particularly at the molecular level.

What should we do with these students? There's much work to be done. But before placing stronger emphasis on aspects of chemistry that we know most students can't handle, perhaps we should begin by facilitating student development of logical reasoning skills through cognitive enrichment experiences prior to their enrollment in the course.

Acknowledgments

The author expresses her gratitude to: Dr. Vantipa Roadrangka and Dr. Michael J. Padilla for granting her permission to translate and administer the GALT test in Spanish at UPR-RP; NIH-PreMARC Grant 5T34GM07821 for partial funding of this project; Statistics Professor Dr. Pedro Rodríguez-Esquerdo for his helpful recommendations; all UPR-RP professors and students who collaborated in this study; and to the reviewers for their useful comments.

References

1. McKinnon JW, Renner J. *Amer. J. Phys* 1971;39:1047–1052.
2. Herron JD. *J. Chem.Educ* 1975;52:146–150.
3. Williams H, Turner CW, Debreuil L, Fast J, Berestiansky J. *J.Chem.Educ* 1979;56:599–600.
4. Niaz M, Robinson WR. *J.Res.Sci.Teach* 1992;29:211–226.
5. Bunce DM, Hutchinson KD. *J.Chem.Educ* 1993;70:183–187.
6. Baird WE, Shaw EL Jr, McLarty P. *Sch.Sci.Math.* February;1996
7. Norman O. *J. Res.Sci.Teach* 1997;34:1067–1081.
8. Ryder, RM.; Pang, Y. The Proceedings of ISECON 2000; §917
9. Anderson, JR.; Reder, LM.; Simon, HA. Applications and Misapplications of Cognitive Psychology to Mathematics Education. [accessed Jan 2009]. 2000 Summer. Texas Educational Review <http://act-r.psy.cmu.edu/papers/146/Applic.MisApp.pdf>
10. Nicoll G, Francisco JS. *J.Chem.Educ* 2001;78:99–102.
11. Samarapungavan A, Robinson WR. *J. Chem.Educ* 2001;78:1107.
12. Bunce DM. *J. Chem.Educ* 2001;78:1107.
13. Legg MJ, Legg JC, Greenbowe TJ. *J.Chem. Educ* 2001;78:1117–1121.
14. Shibley IA Jr, Milakofsky L, Bender DS, Patterson HO. *J.Chem.Educ* 2003;80:569–573.
15. Tai RH, Sadler PM, Loehr JF. *J.Res.Sci.Teach* 2005;42:987–1012.
16. McConnell DA, Steer DN, Owens KD, Knight CC. *J. Geosc.Educ* 2005;53:462–470.
17. Tsapalis G. *Res.Sci.Technol.Educ* 2005;23:125–148.
18. Yaman S. *J. Turk. Sci. Educ* 2005;2:31–33.
19. Lewis SE, Lewis JE. *Chem.Educ.Res.Pract* 2007;8:32–51.
20. Hurst MO, Howard R. *Chem.Educator* 2008;13:42–45.
21. Lawson, AE. Research on the Acquisition of Science Knowledge: Epistemological Foundations of Cognition. In: Gabel, Dorothy L., editor. *Handbook of Research on Science Teaching and Learning: A Project of the National Science Teachers Association*. MacMillan Library Reference USA; New York: 1994. p. 131-176. Chapter 4
22. Roadranga, V.; Yeany, RH.; Padilla, MJ. University of Georgia; 1982.
23. Bunce DM, VanderPlas JR. *Chem.Educ.Res.Pract* 2006;7:160–169.
24. *Microsoft7 Office Excel*, version 2003. Microsoft Corporation; 2003.
25. *SPSS Base*, version 15.0 for Windows7. SPSS, Inc.; 2006.
26. Nunnally, JC. *Psychometric Theory*. 2nd Edition. McGraw-Hill; New York: 1978. Chapter Seven
27. Cronbach LJ. *Psychometrika* 1951;16:297–334.
28. Roadranga, V. PhD Dissertation. University of Georgia; Athens, Georgia: 1986. The Construction and Validation of the Group Assessment of Logical Thinking (GALT). *Dissertation Abstracts International* 46(9):2650.
29. Bitner BL. *J.Res.Sci.Teach* 1991;28:265–274.
30. Nurrenbern SC, Pickering. *J.Chem. Educ* 1987;64:508–510.
31. Lythcott J. *J.Chem. Educ* 1990;67:248–252.
32. Tobias, S. They're Not Dumb, They're Different: Stalking the Second Tier. Research Corporation; Arizona: 1990.
33. Nakhleh MB. *J.Chem. Educ* 1993;70:52–55.
34. Nakhleh MB, Mitchell RC. *J.Chem. Educ* 1993;70:190–192.
35. Zoller U, Lubezky A, Nakhleh MB, Tessier B, Dori YJ. *J.Chem. Educ* 1995;72:987–989.
36. Cracolice MS, Deming JC, Ehlert B. *J.Chem. Educ* 2008;85:873–878.

Figure 1. Final grade distribution in General Chemistry II by operational level (n=400)

Table 1
Student performance by logical reasoning mode

Logical reasoning mode	Mean score	Standard Deviation
Mass/volume conservation	1.53	0.61
Proportional reasoning	0.88	0.79
Experimental variable control	1.36	0.71
Probability	0.97	0.90
Correlation	0.71	0.68
Combinations	1.58	0.53

Table 2
Student performance in different logical reasoning modes by gender

Logical reasoning mode	Mean score by gender		<i>t</i>	<i>Sig.</i> (two-tailed)
	Male	Female		
Mass/volume conservation	1.63 (<i>SD</i> = 0.58)	1.47 (<i>SD</i> = 0.62)	2.68	<i>p</i> < .01
Proportional reasoning	1.15 (<i>SD</i> = 0.76)	0.74 (<i>SD</i> = 0.76)	5.51	<i>p</i> < .001
Experimental variable control	1.33 (<i>SD</i> = 0.71)	1.37 (<i>SD</i> = 0.70)	−0.69	<i>p</i> > .05
Probabilistic reasoning	1.26 (<i>SD</i> = 0.85)	0.83 (<i>SD</i> = 0.88)	5.10	<i>p</i> < .001
Correlational reasoning	0.82 (<i>SD</i> = 0.72)	0.66 (<i>SD</i> = 0.66)	2.37	<i>p</i> < .05
Combinatorial reasoning	1.66 (<i>SD</i> = 0.51)	1.54 (<i>SD</i> = 0.54)	2.39	<i>p</i> < .05

Table 3
Performance in different logical reasoning modes by operational level

Logical reasoning mode	Mean score by operational level			<i>F</i>	<i>Sig.</i>
	Concrete	Transitional	Formal		
Mass/volume conservation	0.97 (<i>SD</i> = 0.56)	1.49 (<i>SD</i> = 0.61)	1.82 (<i>SD</i> = 0.40)	79.93	<i>p</i> < .001
Proportional reasoning	0.16 (<i>SD</i> = 0.37)	0.61 (<i>SD</i> = 0.62)	1.46 (<i>SD</i> = 0.66)	171.76	<i>p</i> < .001
Experimental variable control	0.78 (<i>SD</i> = 0.69)	1.26 (<i>SD</i> = 0.68)	1.72 (<i>SD</i> = 0.50)	74.39	<i>p</i> < .001
Probabilistic reasoning	0.20 (<i>SD</i> = 0.51)	0.60 (<i>SD</i> = 0.78)	1.68 (<i>SD</i> = 0.58)	203.09	<i>p</i> < .001
Correlational reasoning	0.20 (<i>SD</i> = 0.41)	0.55 (<i>SD</i> = 0.59)	1.10 (<i>SD</i> = 0.66)	80.83	<i>p</i> < .001
Combinatorial reasoning	1.14 (<i>SD</i> = 0.51)	1.53 (<i>SD</i> = 0.53)	1.84 (<i>SD</i> = 0.37)	70.51	<i>p</i> < .001

Table 4
Correlation coefficients for each logical reasoning mode as predictor of student performance in General Chemistry (N = 466, df = 464)

Logical reasoning mode	General Chemistry I		General Chemistry II	
	Pearson's R	Sig. (two-tailed)	Pearson's R	Sig. (two-tailed)
Mass/volume conservation	0.125	$p < .005$	0.130	$p < .005$
Proportional reasoning	0.211	$p < .001$	0.172	$p < .001$
Experimental variable control	0.213	$p < .001$	0.199	$p < .001$
Probabilistic reasoning	0.301	$p < .001$	0.288	$p < .001$
Correlational reasoning	0.164	$p < .001$	0.168	$p < .001$
Combinatorial reasoning	0.169	$p < .001$	0.160	$p < .001$

Table 5
Operational level and algorithmic vs. conceptual approach towards a partial exam question (N = 106)

Approach	Operational level			Total
	Concrete	Transitional	Formal	
Algorithmic	10	13	10	33
Algorithmic/Conceptual	11	21	13	45
Conceptual	2	8	18	28

Table 6
Performance in the ACS General Chemistry Examination by operational level

	Mean score by operational level			<i>F</i>	<i>Sig.</i> <i>p</i> < .01
	Concrete	Transitional	Formal		
ACS General Chemistry Examination	27.60 (<i>SD</i> = 6.35)	31.35 (<i>SD</i> = 7.44)	34.50 (<i>SD</i> = 8.93)	17.99	<i>p</i> < .01