Journal of Research in Science Teaching, **1987**, *24*, 645-657

A STUDY OF TWO MEASURES OF SPATIAL ABILITY AS PREDICTORS OF SUCCESS IN DIFFERENT LEVELS OF GENERAL CHEMISTRY

CAROLYN S. CARTER, MARY A. LaRUSSA and GEORGE M. BODNER
Department of Chemistry, Purdue University, West Lafayette, Indiana 47907

Abstract

Preliminary data (Bodner and McMillen, 1986) suggested a correlation between spatial ability and performance in a general chemistry course for science and engineering majors. This correlation was seen not only on highly spatial tasks such as predicting the structures of ionic solids ® = 0.29), but also on tasks such as multiple-choice stoichiometry questions ® = 0.32) that might not be expected to involve spatial skills. To further investigate the relationship between spatial ability and performance in introductory chemistry courses, two spatial tests were given to 1648 students in a course for science and engineering majors (Carter, 1984) and 850 students in a course for students from nursing and agriculture (LaRussa, 1985) at Purdue. Scores on the spatial tests consistently contributed a small but significant amount to success on measures of performance in chemistry. Correlations were largest, however, for subscores that grouped questions that tested problem solving skills rather than rote memory or the application of simple algorithms, and correlations were also large for verbally complex questions that required the students to disembed and restructure relevant information.

Introduction

Studies of spatial ability trace back to the 1920s, when a "practical" or "mechanical" aptitude separate from Spearman's general intelligence factor was first proposed (Smith, 1964). The multiplicity of spatial factors that resulted from early factor analysis studies (Slater, 1940; Guilford and Zimmerman, 1947; Guilford and Lacy, 1947; Thurstone, 1950; French, 1951; Guilford, Fruchter, and Zimmerman, 1952) was eventually reduced to two major factors: *spatial orientation* and *spatial visualization* (Michael, Guilford, Fruchter and Zimmerman, 1957). Spatial orientation has been described as the ability to remain unconfused by changing orientations in which visual stimuli are presented, while spatial visualization involves the ability to mentally manipulate pictorially presented stimuli by a process which involves recognizing, retaining and recalling configurations in which there is movement of the figure or parts of the figure (McGee, 1979).

Experiments on the perception of the upright led Witkin and Asch (1948) to propose a field-dependence/field-independence (FD/FI) construct which was related to an individual's tendency to rely on either the body or the visual field for cues to determining the upright. Witkin and Coworkers hypothesized that the ability to disembed information from a field and then restructure this information was inherent to field independence, and therefore used tests of disembedding in the spatial domain as one measure of FD/FI (Witkin, 1949a and b; Witkin, et al., 1954; Witkin, et al., 1962; Witkin and Goodenough, 1977; Witkin, Goodenough, and Oltman, 1979).

Correlations were found so often between measures of spatial ability and the FD/FI construct (Gardner, Jackson, and Messick, 1960; Gorman, 1968; Haynes and Carley, 1970; Gough and Olton, 1972; Vernon, 1972; Sherman, 1974; Hyde, Geiringer, and Yen, 1975; and Satterly, 1976) that Linn and Kyllonen (1981) eventually questioned whether FD/FI tests measure anything other than spatial ability, and concluded that FD/FI tests believed to measure cognitive restructuring or disembedding were in fact tests of spatial ability, and distinct from measures of perception of the upright. Tests of disembedding and restructuring used in this study will therefore be considered to be tests of spatial ability and not the FD/FI construct.

Preliminary data (Bodner and McMillen, 1986) on the relationship between spatial ability and performance in a general chemistry course for science and engineering majors suggested a significant correlation between these two factors. This correlation was seen not only on highly spatial tasks such as predicting the structures of ionic solids ($r = 0.29$), but also on tasks such as multiple-choice stoichiometry questions ($r = 0.32$) that might not be expected to involve spatial skills. Bodner and McMillen postulated that the relationship between spatial ability and problem solving traces back to the stage of the problem solving process Polya (1945) first described as "understanding the problem".

The spatial tests they used measure the students' ability to disembed and restructure information in the spatial domain. "Understanding the problem" requires a similar ability to disembed pertinent information from the statement of the problem, and restructure or transform this information into a problem the student "understands."

This hypothesis is consistent with the results of Pribyl and Bodner (in press) who found that spatial tests could explain up to 15% of the variance in performance in organic chemistry courses on questions which require problem solving skills. It is also supported by the relationship these authors found between spatial ability and students' representations of problems, which have been assumed to reflect the level of understanding of the problem (Greeno, 1977).

To further investigate the relationship between spatial ability and performance in introductory chemistry courses, two measures of spatial skills were given to students in general chemistry courses for science and engineering majors (Carter, 1984) and agriculture and health science majors (LaRussa, 1985) at Purdue, and performance on these spatial tests was correlated with students' scores on chemistry exams and subscores created by grouping similar exam questions.

Methods

Subjects

The subjects included 850 students enrolled in the first semester of a college-level general chemistry course for students in agriculture and health science (CHM 111) and 1648 students enrolled in the first semester of a general chemistry course for science and engineering majors (CHM 115) at Purdue University during the Fall semester of 1983. Students in CHM 115 constitute a fairly select population of college freshmen, with average SAT math and verbal scores for this sample of 567 and 477, respectively. Students in CHM 111 are a less select population, with average SAT math and verbal scores in this sample of 444 and 396, respectively. CHM 115 was divided into two divisions, called CHM 115M and 115T, which were taught by different faculty but used the same texts and covered the same course content. CHM 115M lectures met at 11:30, 2:30 or 3:30 on Mondays and Fridays; CHM 115T lectures met at either 8:30 or 9:30 on Tuesdays and Thursdays. All three courses (111, 115M and 115T) were taught in large lecture sections of approximately 400 students, with students also meeting in groups of no more than. 4 for recitations and labs. The two paper and pencil spatial tests were administered during the first lab session of each course.

Spatial Ability Tests

Two measures of spatial ability were used in this study: the 20-item version of the Purdue Visualization of Rotations test (Bodner, Carter, and Guay, in press), and the 20-item Find-A-Shape-Puzzle (Linn and Kyllonen, 1981; Linn, Pulos, and Gans, 1981). Data on sample size, mean, standard deviation and estimates of reliability for these tests are given in Table I.

ROT was used in this study because it has been shown to be among the measures of spatial visualization least confounded by analytical processing (Guay, McDaniel, and Angelo, 1978; Guay and McDaniel, 1978). FASP was used because it is believed to measure disembedding in the spatial domain (Linn and Kyllonen, 1981).

TABLE I
Means, Standard Deviations, and Reliabilities of the Spatial Tests

CHM 111

Test	N	Mean	Std. Dev.	Reliability
ROT	850	11.66	3.96	.796 (Split-half)
FASP	850	11.70	5.21	.820 (Cronbach's alpha)

CHM 115 M&T

Test	N	Mean	Std. Dev,	Reliability
ROT	1648	13.96	3.79	.818 (Split-half)
FASP	1648	12.71	5.58	.895 (Cronbach's alpha)

Chemistry Performance

Chemistry achievement was measured by hour exams and final exams written by the faculty in charge of each course. All exams had the same format, consisting of between 25 and 51 multiple-choice questions. No attempt was made to influence the exam format, choice of exam items, course structure, course content, or the assignment of students to a given course or a given division of a course. The exams were assumed to be valid measures of chemistry performance. Their reliability can be estimated from split-half coefficients which ranged from 0.74 to 0.81 for seven of the eight exams in CHM 115M and 115T; exam 3 in CHM 115T had a split-half coefficient of only 0.68.

Subscores were created by grouping similar questions from one or more exams. The subscores covered such diverse topics as stoichiometry, gas laws, crystal structure, molecular geometry, descriptive chemistry, acid-base and redox chemistry, atomic structure and periodic properties, and enthalpy calculations. Subscores were also created which grouped highly verbal questions or questions which focused on either general knowledge or chemical content knowledge. The subscores in this study are described in Tables II and III.

Statistical Procedure

ROT and FASP scores were converted to T-scores and a total spatial score (TSPAT) was calculated for each student by adding these scores and dividing by two. Students were classified as either high, medium or low spatial ability on the basis of their ROT, FASP or TSPAT scores. "High spatial ability" students scored at least one-half standard deviation above the mean on a given score, whereas "low spatial ability" students scored at least one-half standard deviation below the mean. Means standard deviations and estimates of test reliability as well as analysis of variance, analysis of covariance, Pearson product-moment correlation coefficients, and Scheffe's test were all calculated using the SPSS program.

TABLE II
Descriptions of Subscores for CHM 111

Sub-score	Description
EXAM 1	First hour exam (30 multiple-choice questions)
EXAM 2	Second hour exam (30 multiple-choice questions)
EXAM 3	Comprehensive final exam (51 multiple-choice questions)
SUB 1	9 Avogadro's number calculation questions from exams 2 and 3
SUB 2	6 balancing chemical equations questions from exams 1, 2 and 3
SUB 3	6 gram-mole calculation questions from exams 2 and 3
SUB 4	8 empirical formula questions from exams 2 and 4
SUB 5	Sum of all "stoichiometry" questions (Sub 1 through Sub 4)

SUB 6	12 factor label questions from exams 1, 2 and 3
SUB 7	10 general knowledge questions from exams 1 and 2
SUB8	14 chemical content knowledge questions from exams 1 and 2
SUB 9	12 verbal questions from exams 1 and 2
SUB 10	10 quantitative gas law questions from exam 3
SUB 11	6 non-quantitative gas law questions from exam 3

TABLE III

Descriptions of Subscores for CHM 115M

Sub-score	Description
EXAM 1	First hour exam (25 multiple-choice questions)
EXAM 2	Second hour exam (30 multiple-choice questions)
EXAM 3	Third hour exam (30 multiple-choice questions)
EXAM 4	Comprehensive final exam (40 multiple-choice questions)
SUB 1	12 stoichiometry questions from exam 1
SUB 2	8 quantitative gas law questions from exams 1 and 2
SUB 3	5 non-quantitative gas laws questions from exam 1
SUB 4	5 empirical formula questions from exams 1 and 4
SUB 5	9 molecular geometry questions from exams 3 and 4
SUB 6	7 crystal structure questions from exams 3 and 4
SUB 7	5 verbal questions from exams 1, 2 and 4
SUB 8	12 enthalpy calculation questions from exams 3 and 4
SUB 9	10 acid-base or redox questions from exam 3
SUB 10	15 descriptive chemistry questions from exam 2
SUB 11	7 quantitative gas law or stoichiometry questions from exam 4
SUB 12	12 atomic structure or periodic properties questions from exams 2 and 4

DESCRIPTIONS OF SUBSCORES FOR CHM 115T

Sub-score	Description
EXAM 1	First hour exam (25 multiple-choice questions)
EXAM 2	Second hour exam (32 multiple-choice questions)
EXAM 3	Third hour exam (39 multiple-choice questions)
EXAM 4	Comprehensive final exam (40 multiple-choice questions)
SUB 1	12 stoichiometry questions from exam 1
SUB 2	6 quantitative gas law questions from exam 1
SUB 3	4 non-quantitative gas law questions from exam 1
SUB 4	12 atomic structure or periodic properties questions from exam 2
SUB 5	12 molecular geometry questions from exam 2
SUB 6	11 crystal structure questions from exams 3 and 4
SUB 7	15 true/false questions from exam 3
SUB 8	7 enthalpy calculation questions from exams 3 and 4

SUB 9	6 acid-base or redox questions from exam 3
SUB 10	13 descriptive chemistry questions from exam 4
SUB 11	6 quantitative gas law or stoichiometry questions from exam 4
SUB 12	9 atomic structure and periodic properties questions from exam 4

Results

Analysis of variance showed a statistically significant difference in performance on all of the exams and 24 of the 35 subscores when students were classified as high, medium or low spatial ability on the basis of ROT scores; on all exams and 30 of the 35 subscores when students were classified on the basis of FASP scores; and on all exams and 32 of the 35 subscores when students were classified on the basis of the total spatial score (TSPAT). Results of the analysis of variance for TSPAT scores in CHM 111, 11SM and 11ST are given in Table IV. Scheffe's test (Scheffe, 1953) was used to determine the direction of difference in chemistry achievement. Results of this test showed that students classified as "high spatial ability" on TSPAT significantly outperformed "low spatial ability" students on all of the exams, 10 of the 11 CHM 111 subscores, and 16 of the 24 CHM 115 subscores.

SAT mathematics and verbal scores were used as covariates in CHM 115M and 115T to determine whether the spatial tests measured a factor beyond general math and verbal skills or test-taking ability. Using these factors as covariates slightly reduced F ratios between spatial ability and chemistry achievement, but in 19 out of 32 cases spatial scores were still statistically significant. In another 3 cases the correlation between SAT and spatial scores was too large to allow analysis of covariance to be done. Thus, when the skills measured by the SAT exams are taken into account, spatial ability is still a factor in these courses. In CHM 111, correlations between SAT scores and the spatial tests were too large to allow analysis of covariance to be done.

TABLE IV
Values from Analysis of Variance for TSPAT in CHM 111, 115M, and 115T

Sub-score	*Source*	CHM 111	CHM 115M	CHM 115T
Exam 1	TSPAT	16.70 (df = 775)***	22.08 (df = 488)***	11.86 (df = 630)***
	SEX	0.34	9.80 **	0.92
Exam 2	TSPAT	19.64 (df = 766)***	13.90 (df = 448)***	4.63 (df = 574)*
	SEX	1.60	14.67 ***	1.86
Exam 3	TSPAT	27.68 (df = 753)***	13.90 (df = 442)***	9.54 (df = 568)***
	SEX	2.35	9.79 **	0.80
Exam 4	TSPAT		5.49 (df = 442)**	13.36 (df = 560)***
	SEX		11.96 ***	1.30
Sub 1	TSPAT	12.95 (df = 740)***	6.04 (df = 441)**	12.64 (df = 630)***
	SEX	16.94 ***	8.47 **	1.63

Sub 2	TSPAT	14.15 (df = 740)***	12.43 (df = 441***	11.38 (df = 630)***
	SEX	0.74	22.49 ***	0.24
Sub 3	TSPAT	12.94 (df = 740)***	14.00 (df = 441)***	3.67 (df = 630)*
	SEX	0.00	8.65 **	0.03
Sub 4	TSPAT	14.35 (df = 740)***	16.63 (df = 441)***	4.44 (df = 630)*
	SEX	11.79 ***	10.04 **	0.96
Sub 5	TSPAT	14.35 (df = 740)***	4.67 (df = 441)**	2.75 (df = 574)
	SEX	15.83 ***	3.98 *	0.02
Sub 6	TSPAT	33.88 (df = 740)***	12.87 (df = 441)***	10.28 (df = 574)***
	SEX	17.51 ***	6.59 *	0.85
Sub 7	TSPAT	2.13 (df = 740)	8.19 (df = 488)***	3.63 (df = 556)*
	SEX	5.39 *	18.35 ***	0.19
Sub 8	TSPAT	7.10 (df = 740)***	6.05 (df = 488)**	7.99 (df = 556)***
	SEX	0.04	14.16	*** 0.88
Sub 9	TSPAT	13.74 (df = 740)***	5.85 (df = 488)**	2.40 (df = 556)
	SEX	1.81	2.97	0.01
Sub 10	TSPAT	19.21 (df = 740)***	5.06 (df = 488)**	4.43 (df = 556)*
	SEX	1.56	7.61 **	0.45
Sub 11	TSPAT	8.97 (df = 740)***	11.91 (df = 488)***	10.64 (df = 556)***
	SEX	0.06	7.82 **	1.85
Sub 12	TSPAT		11.84 (df = 488)***	9.18 (df = 556)***
	SEX		4.26 *	0.66

*p < 0.05 **p < 0.01 ***p < 0.001

Sex was used as an independent variable in this study because preliminary work (McMillen, 1983) suggested that males in a similar population tend to score higher than females on the ROT test. Sex was a significant contributor to virtually every exam and subscore in CHM 115M, but made no significant contribution in CHM 115T. The presence of sex as a significant factor in 115M but not 115T is impossible to explain because it could result from so many factors. It might result from differences between the professors who taught the two courses, and the degree to which they emphasized the quantitative or mathematical aspects of chemistry versus descriptive or qualitative concepts. It is also likely to reflect differences between the exam questions use in the two courses, as well as anisotropy in the distribution of students between courses. Students are not randomly assigned to these courses; whether they are scheduled to attend chemistry lectures on M and F or T and Th depends on the pattern of other courses for which they register. In CHM 111, sex was a significant contributor to subscores which focused on factor-label or stoichiometry calculations. No significant interactions between sex and spatial score were found on any measures of chemistry achievement in either 115M or 115T, while significant interactions were observed in 111 only on subscores which focused on factor-label or stoichiometry calculations.

Pearson product-moment correlation coefficients (Tables V-VI) show small but highly significant correlations between spatial ability and achievement on most scores.

Correlations tend to be higher on subscores believed to be measures of problem solving skills than subscores which feature questions that can be answered algorithmically or from memory. Multiple regression analysis using ROT scores, FASP scores and sex as independent variables produced total correlations with exams and subscores ranging from 0.10 to 0.37.

T-tests were used to judge differences between students who dropped CHM 115M or 115T and those who completed these courses, as well as differences between students who received A or B grades in these courses and those who received D or F grades. No significant difference on either ROT, FASP or TSPAT was found between students who dropped 115M or 11ST and those who completed the course. On the other hand, statistically significant differences (p <0.0001) were found between A/B and D/F students in CHM 115M or 115T or ROT, FASP and TSPAT scores.

TABLE V
Correlations between Total Spatial Scores and Chemistry Sub-Scores

	CHM 111		
Sub-score	ROT	FASP	TSPAT
EXAM 1	.20 *	.18 *	.20 *
EXAM 2	.17 *	.19 *	.23 *
EXAM 3	.25 *	.24 *	.23 *
SUB 1	.20 *	.16 *	.20 *
SUB 2	.19*	.15 *	.20 *
SUB 3	.18 *	.13 *	.18 *
SUB 4	.16 *	.16 *	.21 *
SUB 5	.24 *	.20 *	.26 *
SUB 6	.29 *	.26 *	.32 *
SUB 7	.04	.07	.05
SUB 8	.13 *	.12 *	.14 *
SUB 9	.20 *	.15 *	.20 *
SUB 10	.21 *	.19 *	.24 *
SUB 11	.15 *	.14 *	.16 *

*p < 0.001

TABLE VI
Correlations between Total Spatial Scores and Chemistry Sub-Scores

	CHM 115M			CHM II5T		
Sub-score	ROT	FASP	TSPAT	ROT	FASP	TSPAT
EXAM 1	.25 **	.23 **	.30 **	.16 **	.17 **	.18 **
EXAM 2	.16 **	.22 **	.23 **	.13 **	.07	.11 *
EXAM 3	.17 **	.21 **	.22 **	.15 **	.12 *	.17 **

EXAM 4	.17 **	.18 **	.19 **	.19 **	.17 **	.21 **
SUB 1	.09	.13 *	.11 *	.17 **	.14 **	.19 **
SUB 2	.22 **	.20 **	.27 **	.14 **	.15 **	.18 **
SUB 3	.22 **	.17 **	.24 **	.05	.09	.09
SUB 4	.23 **	.23 **	.25 **	.14 **	.05	.12 *
SUB 5	.07	.13 *	.11 *	.12 *	.05	.11 *
SUB 6	.19 **	.17 **	.20 **	.17 **	.15 **	.19 **
SUB 7	.11 *	.16 **	.17 **	.07	.10	.10
SUB 8	.12 *	.15 **	.16 **	.15 **	.11 *	.15 **
SUB 9	.12 *	.17 **	.15 **	.06	.05	.09
SUB 10	.08	.14 **	.13 **	.10 *	.11 *	.12 *
SUB 11	.16 **	.22 **	.19 **	.16 **	.15 **	.17 **
SUB 12	.13 *	.18 **	.20 **	.14 **	.13 **	.16 **

* $p< 0.01$ ** $p < 0.001$

Discussion

Results of the analysis of variance, Pearson product-moment correlation calculations, and multiple regression analysis all suggest that for CHM 111 the total spatial score is most strongly correlated to subscore 6 which included questions which involve unit conversion calculations such as:

> In Apothecaries' measurement, 1 dram = exactly 60 grains and 1 pound = exactly 96 drams. What is the mass in grams of aspirin in a 15.0 grain aspirin tablet? (a) 1.9 x 102 g Co) 2.6 g (c) 1.2 g (d) 0.25 g (e) 5.7 X $10^{'6}$ g

or:

> A 19.5 g cube of copper (density = 8.92 g/mL) was placed on the bottom of a graduated cylinder. What volume of ethyl alcohol (density = 0.789 g/mL) should be added in order to raise the meniscus to the 25.0 mL mark? (a) 5.5 mL (b) 16.1 mL (c) 18.0 mL (d) 22.8 mL (e) none of these

Although these questions can be answered using the factor-label algorithm, they require more than just a mindless application of this algorithm. In one case there appears to be too much information, in the other case there appears to be not enough. For students in CHM 111, these questions might fit better into the category of "problems" rather than "exercises" as these terms were defined by Bodner and McMillen (1986).

TSPAT also correlated well with subscore 10 in CHM 111 which included quantitative gas law questions such as:

> 1.00 L of nitrogen gas, initially at 30.0°C and a pressure of 1.60 arm, is compressed to a volume of 0.750 L while its pressure is increased to 2.00 atm. Calculate the

new temperature of the gas in °Celsius. (a) 284°C (b) 232°C (c) 28°C (d) 11°C (e) -91 °C

and:

Strontium sulfate decomposes upon heating:
$SrSO_4(s) \rightarrow SO_3(g) + SrO(s)$

How many liters of SO_3 at STP can be produced by the complete decomposition of 177 g $SrSO_4$? (a) 7.21 L (b) 10.8 L (c) 21.6 L (d) 23.2 L (e) 43.2 L

These questions are often solved algorithmically by general chemistry instructors, but they are more likely to be viewed as problems by beginning students. TSPAT also correlated well with subscore 4 in CHM 111 which included empirical formula questions that can be solved algorithmically by beginning students who are drilled in this type of calculation, such as:

The first chemical compound of a noble gas element was prepared in 1962. What is the empirical formula of a compound of xenon and oxygen which is 67.2% Xe and 32.8% O? (a) XeO_2 (b) XeO_5 (c) XeO_3 (d) XeO_6 (e) XeO_4

But this subscore also included questions which are much less likely to be solved algorithmically, such as:

9.33 grams of copper metal was allowed to react with an excess of chlorine and it was found that 14.6 grams of a compound of copper and chlorine were formed. What is the empirical formula of this compound? (a) Cu_2Cl (b) $CuCl_2$ (c) $CuCl_3$ (d) CuCl (e) $CuCl_5$

The correlations were much smaller in CHM 111 for subscore 7 which included general knowledge questions such as:

Which of the following is not one of the base units in the International System? (a) ampere (b) liter (c) kilogram (d) Kelvin (e) second

and:

Which of the following metric prefixes is incorrectly matured with its numerical value? (a) mega, 10^6 (b) deci, 10^{-1} (c) kilo, 10^3 (d) micro, 10^{-6} (e) milli, 10^{-3}

Correlations were also smaller for subscore 8 which included chemical content knowledge questions that can be answered from memory or by application of well-established algorithms, such as:

The electron configuration: $ls^2\,2s^2\,2p^6\,3s^2\,3p^6$ represents which species? (a) Ar (b)

Cl (c) K (d) Mg (e) no correct response

or:

Which of the following compounds is incorrectly named?

(a) $CoBr_2$ – cobalt(II) bromide
(B) AgCN – silver(III) cyanide
(c) $V_2(SO_4)_3$ – vanadium(III) sulfate
(d) $Mn_3(PO_4)_2$ – manganese(II) phosphate
(e) $CuCl_2$ – copper(II) chloride

Similar results can be obtained by analyzing the CHM 115 subscores. Correlations are largest for questions which are most likely to involve problem solving, such as those in subscore 4 of CHM 115M:

> Uranium reacts with fluorine to produce a compound which is a gas at 57°C. The density of this gas is 13.0 g/L at 57°C and 1 atm pressure. What is the molecular formula of this compound? (a) UF_2 (b) UF_3 (c) UF_4 (d) UF_5 (e) UF_6

Correlations tend to be smallest for questions which can be answered algorithmically or from memory, such as those in subscore 4 of CHM 115T:

> Which of the following correctly lists the elements in increasing order of electronegativity?
>
> (a) Sb < As < Te < At
> (b) Sb < As < Se < Br
> (c) Br < Se < As < Sb
> (d) At < Te < As < Sb
> (c) none of these is correct

Evidence for the role of disembedding and cognitive restructuring in chemistry can be obtained by noting that true-false questions which tested for chemistry content knowledge such as:

> H_3PO_4 is a stronger acid thanH_4SiO_4
> (a) true
> (b) false

in subscore 7 in CHM 115T correlated very poorly with TSPAT, whereas chemical content knowledge questions which required more extensive amounts of disembed-ding and restructuring such as:

> Which of the following statements is correct7

(a) Isotopes of an element have the same mass numbers.
(b) Hydrogen and deuterium are not isotopes because they have different symbols.
(c) Potassium and sodium are considered to be isotopes because they have very similar chemical properties.
(d) Atomic weights are averages of all the isotopic masses of a given element. relative to the mass of a reference isotopic mass.
(e) When ^{22}Mg becomes a Mg^{2+} ion, it is no longer isotopic with ^{24}Mg because it has lost two electrons.

in subscore 9 of CHM 111, or:

Which of the following statements explains why a hot air balloon rises when the air in the balloon is heated?

(a) As the temperature of the gas increases, the average kinetic energy of the gas molecules increases, and the collisions between these gas molecules and the walls of the balloon makes the balloon rise.
(b) As the temperature of the gas increases, the pressure of the gas increases, pushing up on the balloon.
(c) As the temperature of the gas increases, the gas expands, some of the gas escapes from the bottom of the balloon, and the decrease in the density of the gas in the balloon lifts the balloon.
(d) As the temperature of the gas increases, the volume of the balloon expands, causing the balloon to rise.
(e) As the temperature of the gas increases, the hot air rises inside the balloon, and this produces enough force to lift the balloon.

in subscore 7 of CHM 115M correlate much more positively with TSPAT.

Conclusions

Neither this study nor our previous work can be viewed as conclusive evidence for the importance of an early stage in problem solving in which relevant information is disembedded from a question and the question is transformed or restructured into a problem for which the student understands the initial and final or goal states. However, the consistency with which we have found correlations between tests of disembedding and restructuring in the spatial domain and performance in chemistry on tasks which require problem solving skills rather than rote memory or the application of simple algorithms might be considered to support this hypothesis.

More importantly, these studies remind chemists that the stage known as "understanding the problem" is the essence of problem solving. If we define problem solving as "what we do when we don't know what to do" (Wheatley, 1984), it is tempting to suggest that by the time we reach the point where we "understand" a problem, the problem solving process is over. These studies may also explain why efforts to teach problem solving which focus exclusively on the analytic processes used to obtain answers to questions seldom take students beyond the point where they can solve familiar exercises. Finally, these studies encourage us to find ways to help students successfully complete the stage of problem solving in which they build an understanding of the problem they have been asked to solve.

References

Bodner, G.M., Carter, C. S., & Guay, R. (in press). The Purdue visualization of rotations test. *Journal of Research in Science Teaching.*

Bodner, G.M., & McMillen, T. L. B. (in press). Cognitive restructuring as an early step in problem solving. *Journal of Research in Science Teaching.*

Carter, C. S. (1984). *A study of the relationships between spatial ability, Sex and success in general chemistry.* Unpublished master's thesis, Purdue University, West Lafayette, IN.

French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monograph,* No. 5.

Gardner, R. W., Jackson, D. N., & Messick, S. J. (1960). Personality organization in cognitive controls and intellectual abilities. *Psychological Issues,* 2, No. 4, Monograph 8.

Gorman, B. S. (1968). Field dependence and visual maze learning. *Perceptual and Motor Skills,* 27, 142.

Gough, H. G., & Olton, R. M. (1972). Field independence as related to non-verbal measures of perceptual performance and cognitive ability. *Journal of Consulting and Clinical Psychology,* 38,338-342.

Greeno, J. G. (1977). Processes of understanding in problem solving. In *Cognitive theory,* Vol 2, N.J. Castellan, D. B. Pisoni, and G. R. Potts (Eds.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Guay, R., McDaniel, E., & Angelo, S. (1978). *Analytic factor confounding spatial ability measurement.* Paper presented at the meeting of the American Psychological Association, Toronto, Ontario, Canada.

Guay, R. B. & McDaniel, E. (1978). *Correlates of performance on spatial aptitude tests.* Final Report for the U.S. Army Research Institute.

Guilford, J.P. & Lacey, J. I. (1947). *Printed classification tests, A.A.F.* Aviation Psychological Progress Research Report, No. 5, Washington, D.C., U.S. Government Printing Office.

Guilford, J.P. & Zimmerman, W. S. (1947). Some A.A.F. findings concerning aptitude factors. *Occupations,* 26,154-159.

Guilford, J.P., Fruchter, B., & Zimmerman, W. S. (1952). Factor analysis of the Army Air Forces, Sheppard field battery of experimental aptitude tests. *Psychometrika,* 16, 45-68.

Haynes, J. R., & Carley, J. W. (1970). Relation of spatial abilities and selected personality traits. *Psychological Reports,* 26, 214.

Hyde, J. S., Geiringer, E. R., & Yen, W. M. (1975). On the empirical relation between spatial ability and sex differences in other aspects of cognitive performance. *Multivariate Behavioral Research,* 10,289-310.

LaRussa, M. A. (1985). *Spatial ability, hemispheric specialization and achievement in chemistry.* Unpublished masters dissertation, Purdue University, West Lafayette, IN.

Linn, M. C. & Kyllonen, P. (1981). The field dependence-independence construct: Some, one, or none. *Journal of Educational Psychology,* 73,261-273.

Linn, M. C, Pulos, S. & Gans, A. (1981). Correlates of formal reasoning: Content and problem effects. *Journal of Research in Science Teaching,* 18,435-447.

McGee, M. G. (1979). *Human spatial abilities: Sources of sex differences.* New York: Praeger.

Michael, W. B., Guilford, $. P., Fruchter, B. & Zimmerman, W. S. (1957). The description of spatial-visualization abilities. *Educational and Psychological Measurement,* 17, 185-199.

Polya, G. (1945). *How to solve it: A new aspect of mathematical method.* Princeton, NJ: Princeton University Press.

Pribyl, J. R. & Bodner, G. M. (in press). Spatial ability and its role in organic chemistry: A study of four courses. *Journal of Research in Science Teaching.*

Satterly, D. J. (1976). Cognitive styles, spatial ability, and school achievement. *Journal of Educational Psychology,* 68, 36-42.

Scheffe, H. (1953). A method for judging all contrasts in the analysis of variance. *Biometrics,* volume XL, June.

Sherman, J. A. (1974). Field articulation, sex, spatial visualization, dependency, practice, laterality of the brain and birth order. *Perceptual and Motor Skills,* 38, 1223-1235.

Slater, P. (1940). Some group tests of spatial judgement or practical ability. *Occupational Psychology,* 14, 40-55.

Smith, I. M. (1964). *Spatial ability, its educational and social significance.* San Diego, CA: Robert R. Knapp.

Thurstone, L. L. (1950). *Some primary abilities in visual thinking.* Chicago, IL: University of Chicago Psychometric Lab Report No. 59.

Vernon, P. E. (1972). The distinctiveness of field dependence. *Journal of Personality,* 40,366-391.

Wheatley, G. H. (1984). *Problem solving in school mathematics.* MEPS Technical Report 84.01, School Mathematics and Science Center, Purdue University, West Lafayette, IN.

Witkin, H. A. & Asch, S. E. (1948). Studies in space orientation. I-IV. *Journal of Experimental Psychology,* 38,325-337; 455-477; 603-614; and 762-782.

Witkin, H. A. (1949a). Perception of the body position and of the position of the Visual Field. *Psychological Monograph,* No. 302, 63, No. 7.

Witkin, H. A. (1949b). The nature and importance of individual differences in Perception. *Journal of Personality,* 18, 145-170.

Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissener, P. B., & Wapner, S. (1954). *Personality through perception.* New York: Harper.

Witkin, H. A., Dyk, R. B., Fatarson, G. E., Goodenough, D. R., & Karp, S. A. (1962). *Psychological differentiation.* New York: Wiley.

Witkin, H. A. & Goodenough, D. R. (1977). *Field dependence revisited* Research Bulletin 77-117, Princeton, NJ: Educational Testing Service.

Witkin, H. A., Goodenough, D. R., and Oltman, P. K. (1979). Field dependence and interpersonal behavior. *Psychological Bulletin,* 84,661--689.