Cognitive factors that can potentially affect pupils' test performance

Eleni Danili and Norman Reid

Centre for Science Education, University of Glasgow, Glasgow G12 8QQ, UK
e-mail: n.reid@mis.gla.ac.uk

Received 23 August 2005, accepted 9 February 2006

Abstract. The two cognitive styles, 'field dependent/field independent' and 'convergent/divergent', were explored in relation to three formats of assessment (multiple choice, short answer and structural communication grid) in five classroom chemistry tests. The study was conducted in Greece with the participation of Grade-10 (upper secondary school) pupils (age 15-16). The field dependent/independent characteristic correlated with pupils' performance in all the tests, and in almost all the formats of assessment. The convergent/divergent characteristic correlated with pupils' performance in assessment where language was an important factor, but not in algorithmic types of questions or in questions where there is a greater use of symbols and less use of words. It seems that, in relation to the convergent/divergent characteristic, the chemistry content and presentation of the test is a factor affecting the type of questions being asked. This study suggests that some of the factors that affect pupils' performance might be: (a) the content and presentation of the test, (b) the format of the test, (c) the psychology of the individual. [*Chem. Educ. Res. Pract.,* 2006, **7** (2), 64-83]

Keywords: assessment in chemistry, cognitive factors, cognitive styles, field dependence/independence, convergence/divergence

Background

In a previous study, correlations between different formats of assessment were found to be between 0.30 and 0.71 (Danili and Reid, 2005). This is a wide range and even the maximum of the correlations is less than 1 by a significant margin. This suggests that the best student found by one method is not necessarily the best student by another method. If the two formats of assessment were simply testing the same content, then very high correlation would be expected. Of course it can be argued that different formats of assessment test different abilities of the examinees and, therefore, it is fairer to use several formats to assess student skills and knowledge. However, the fundamental issues arising from the above study were:

- Are the different formats of questions testing different abilities? Which format of assessment is more valid?
- Are the different formats related to differences between students in one or more psychological traits?
- It might be reasonable to suppose that the use of multiple formats of assessment tests students more fairly than the use of a single format but on what basis can it be justified?

To address some of these questions, a study was designed which engaged 476 pupils (Grade 10, age 15-16) of 12 public upper secondary schools in Greece during the school year September 2002-May 2003. In this study, the two cognitive styles ‘field dependent/field independent’ and ‘convergent/divergent’ were explored in relation to three formats of assessment (multiple choice, short answer and structural communication grid) in five classroom chemistry tests.

Assessment formats in a dynamic interaction with teaching and learning models

Teaching and assessment are inseparable in the learning process. Assessment does not stand outside teaching and learning but stands in a dynamic interaction with them. Shepard (1992) emphasised the importance for educators to understand the conception of teaching and learning when they make decisions about testing practice and to examine the implicit theories which guide their practice. In the traditional model of teaching (sometimes described as objectivism as distinct from constructivism) learning is seen in terms of a distinct body of information, specified in detail, that can be transmitted to the learner. Assessment, in this context, consists of checking whether the information has been received (Entwistle, 1992). However, isolated facts, if learnt, are often not easily recalled from the memory because they have no meaning and do not fit into the learner’s conceptual map. Students can succeed in objective tests without necessarily understanding the material they have learned. This may be true particularly in science where much research has shown that students carry widespread misconceptions and misunderstanding of scientific phenomena (Osborne and Cosgrove, 1983; Nurrenbern and Pickering, 1987; Sawrey, 1990; Andersson, 1990; Bodner, 1991; Gabel, 1999). The behaviourist learning theory requires practice, repetition and testing of discrete basic skills prior to any teaching of higher-order thinking skills (Shepard, 1992).

In contrast, in the constructivist and information processing models learning is seen as a process of personal knowledge construction and meaning making. In this approach, learning is a complex and diverse process and therefore requires assessment to be more diverse, and to assess in more depth the structure and quality of students’ learning and understanding (Gipps, 1994). In the information processing models, the structure of effective learning is seen in such a way that knowledge can be stored usefully in the long-term memory. Knowledge is seen as something cohesive and holistic which provides scaffolding for later learning (Atkins et al., 1992). In fact, cognitive processes indicate that there is an intimate connection between skills and the contexts in which they are used. This means that assessment should reduce the emphasis on the ability to memorise, and increase the emphasis on thinking and problem solving. Information processing approaches to learning require a new assessment methodology, and tests ought not to ask for demonstration of small, discrete skills practised in isolation (Gipps, 1994).

The importance of aligning teaching methods and assessment tasks is stressed in many publications pertaining to the curriculum (Osborne, 2004). However, over the last decade, the amount of assessment in schools has increased. Consequently, the assessment workload for the teachers has grown dramatically, and the time available to devote to assessing each student has fallen. It is tempting to reduce marking workload by using objective tests. Objective testing assessment policy is based on objectivistic theories and is greatly concerned with quantitative measurement (Biggs, 1996). The quality of such assessment is embodied in notions of reliability and validity (Broadfoot and Black, 2004). Unfortunately, objective assessment practices can inadvertently de-skill students in various ways. They focus attention on the immediate tasks of passing examinations or completing tasks and distract students from the more vital task of learning how to assess themselves (Boud, 2004). This tradition is very much opposite to a *constructivist* theory of learning, which regards learning more in qualitative than quantitative terms (Biggs, 1996). According to constructivist theory, assessments policy should be based on performance on open-ended tasks which can reveal a wide variety of insights of thinking processes in students' written responses.

Cognitive styles

Learning theories are the bases which help teachers and educators to understand diverse factors of individual differentiation in: perceiving information; encoding information; transferring information; scanning the representation of the information; and working memory capacity. There are also individual differences in styles of remembering, thinking, and judging, and these individual variations, if not directly part of the personality, are at the very least intimately associated with various non-cognitive dimensions of personality (Kogan, 1976). Differences in the above factors are brought together to suggest that individuals have different cognitive styles and are different in intelligence, ability, personality, and achievement. It seems that our cognitive style influences our: intellectual abilities; skills; personalities; teaching and learning; and performance. According to Messick's (1993) definition "*cognitive styles are characteristic modes of perceiving, remembering, thinking, problem solving, decision making that are reflective of information processing regularities that develop in congenial ways*".

There have been various arguments relating to the overlap between style and ability. Some researchers support the idea that 'ability' describes performance in a given task whereas 'style' describes the way the task is approached (Messick, 1994) While intellectual abilities are primarily concerned with the ability to learn, cognitive styles are primarily concerned with differences in the ways of learning. Riding and Cheema (1991) considered cognitive style to be a fairly fixed characteristic of an individual while cognitive strategies are the ways that may be used to cope with particular situations and tasks. Strategies may be learned and developed. Styles, by contrast, are static and are relatively in-built features of the individual.

In the literature there are various labels of cognitive styles. In this study attention was focussed on field dependent/independent and convergent/divergent cognitive styles. The reasons for that were:

- they are dominant over the other cognitive styles in the literature.
- previous work suggests that they are related to assessment.

Field-dependent /independent cognitive style

Hundreds, if not thousands, of articles pertaining to the field dependence-independence (FDI) construct have been published. This polar construct originated in Witkin's work (Witkin et al., 1974; Witkin et al., 1977; Witkin and Goodenough, 1981). Witkin and Goodenough (1981) investigated for many years the idea suggested by Gestalt psychology, that some people are dominated by any strong frame of reference or pattern in a stimulus field, to such an extent that they have trouble in perceiving elements that cut across the pattern. He investigated the personality in relation to the integrative process of making contact with the environment through perception.

Early studies of Witkin and Asch (1948a, 1948b) found that some individuals consistently tended to attend to different type of cues. Subjects who used visual cues were designated 'field-dependent', while those who used postural cues (such as tactile, vestibular and kinaesthetic cues) were designated 'field-independent'. Further probes of the subject's ability to perceive individual elements within an organised perceptual field have followed. It was thought that there might be a relationship between the individuals 'disembedding ability' and their 'cognitive restructuring'.

Within this framework, Witkin and Goodenough (1981) defined the main characteristic of the field-dependent and field-independent cognitive styles as:

- *Field - Dependent* (FD) individual: one who can insufficiently separate an item from its context and who readily accepts the dominating field or context.
- *Field - Independent* (FID) individual: one who can easily 'break up' an organised perceptual field and separate readily an item from its context.

A number of studies have followed in examining the correlation between field dependency/independency (FDI) and academic performance in disciplines such as language, mathematics, natural sciences, social sciences, art, music and computer science at secondary school level as well as at university level. Tinajero and Paramo's (1998) review concluded that "in general field-independent subjects perform better than field-dependent subjects, whether assessment is of specific disciplines or across the board".

Many research studies (e.g. Johnstone and El-Banna, 1986; Al-Naeme and Johnstone, 1991; Bahar and Hansell, 2000; Danili and Reid, 2004; Tsaparlis, 2005) have looked at FDI. Their findings are that students' performances are consistent with the conclusion of Tinajero and Paramo (Tinajero and Paramo, 1998). Overall, the field dependent/independent test is considered by many researchers a very powerful instrument to predict academic performance of individuals (e.g. Terrell, 2002).

Convergent / divergent cognitive style

Research on the Convergence-Divergence cognitive style has not received as much attention as the FDI cognitive style from educators and researchers. The idea of convergent-divergent cognitive style has its origin with Hudson (1966) who, as an undergraduate, had found himself better at some parts of intelligence tests than others: good at the diagrammatic questions, and relatively poor at the verbal and numerical ones. At that time, there was a growing feeling that typical intelligence tests did not measure all aspects of intelligence. It was argued that such tests only measured what was termed 'convergent thinking' and not 'divergent thinking'. Convergent thinking means that someone has to focus down (converge) on the one right answer in order to find the solution of a problem. Convergent thinkers score highly in problems requiring one conventionally accepted solution clearly obtainable from the information available (as in intelligence tests), while at the same time obtaining low scores in problems requiring the generation of several equally acceptable solutions. On the other hand, divergent thinking is the opposite of this approach. Divergent thinking deals with the capacity to generate responses, to invent new ones, to explore and expand ideas, and in a word, to diverge. Convergent thinking thus demands close reasoning; divergent thinking demands fluency and flexibility (Child and Smithers, 1973).

Hudson (1966, 1968) thought that he might be able to measure arts/science aptitude and made an attempt to devise tests of aptitude for arts and science respectively. In the traditional IQ test, the individual is required to find the one right answer for a problem, being invited to choose this right answer from a list of alternatives. The new tests do not require the respondent to produce one right answer, and like intelligence tests, can take different forms. Such questions can consider imaginative themes by asking questions such as:

- How many uses can you think of for each of the following objects?
- How many meanings can you think of for each of the following words?

- Draw a picture in the space below to illustrate the title 'Zebra Crossing'.

According to Hudson (1966) "*the converger is the boy (*sic*) who is substantially better at the intelligence test than he is at the open-ended tests; the diverger is the reverse*".

Most of the research related to convergent/divergent styles has concentrated on the relationship between divergent thinking and arts-science orientation. Research showed that there is a tendency for convergers to choose science subjects. Johnstone and Al-Naeme (1995) indicated that much science teaching is convergent and students are rewarded for convergent thinking leading to unique specific answers. However, this may not to be the case for biology because it attracts both groups of students (Orton, 1992; Bahar, 1999). Bahar's statement was that, "*biology might be one of the science branches in which students might cope equally well with a convergent or a divergent bias*".

Many researchers tended to equate divergent thinking with creativity and convergent thinking with intelligence. This has caused a great deal of controversy, with different research supporting different results (Nuttall, 1972; Bennett, 1973; Runco, 1986; Fryer, 1996). In the literature little research is reported on convergent/divergent cognitive styles and performance in science. However, Al-Naeme's (1988) research showed that divergent students had higher scores than convergent students in mini projects in chemistry.

Bahar (1999) showed that divergent pupils/students did not perform better in all cases compared to convergent pupils/students. Thus, he suggested that the answer might be related with assessment techniques. He said "…*when one is looking at the relationship between students' performance in any topic and their cognitive styles, the type of assessment techniques used, such as multiple choice type of questions, essay questions, projects and so forth should be reported because a particular type of assessment technique may favour a particular kind of cognitive style*".

In general, it seems that the language and the format of questions in relation to the cognitive style of an individual may be able to influence his/her performance.

What should be the aims of educational measurement?

From the above, it can be concluded that cognitive styles influence the personality of the individuals, and affect the psychological behaviours that indicate how learners perceive, interact with and respond to the learning environment **(Fatt, 2000).** Accordingly, cognitive styles have an impact on pupils' performance and achievement. Therefore, the concern of educators should be to understand, from the heterogeneous mix of pupils' learning styles, the possible styles, so that teachers can best adapt their teaching style and assessment materials to suit the pupils' preferred styles and help them to overcome their difficulties and display their abilities. This is a daunting prospect for the teacher!

Furthermore, if assessment is to be part of teaching, then first it has to be seen that way. Most areas of learning have both mental and physical aspects. However, all learning has an emotional aspect and numerous research studies emphasise the importance of learner confidence, motivation and self-esteem, which are prerequisites for successful learning, and need to be encouraged. Therefore, the negative or positive impact of different forms of assessment on motivation and self-esteem need to be considered seriously. Thus, there is a need to reinforce pupils' motivation by assessing them with appropriate format questions and, therefore, to enable them to show their best performance (Gipps, 1994). Assessment must be humane (Johnstone, 2003). Humanity takes into account factors that affect pupils/students' performance, such us cognitive and psychological traits of individual personality.

If the purpose of educational measurement is 'how well' rather than 'how many', then this requires a quite different approach to test construction. Gipps (1994) pointed out that we need a more measured, analytical approach to assessment in education. We need to resist the tendency to think in simplistic terms about one particular form of assessment being better than other: consideration of form without consideration of purpose is wasted effort. She called for wider understanding of the effect of assessment on teaching and learning and fostering a system, which supports multiple methods of assessment while at the same time making sure that each one is used appropriately.

Thus, an important question in education is: what should be the aims of educational measurement? Should the aim be to devise tests, which look at the individual and find out 'how well' or should look at 'how many'? Should the aim be to devise tests to support learning or to devise tests to evaluate teachers according to pupils' performance and achievement? The dangers of 'teaching to the test' are well known and, if only a limited range of facts and skills are assessed, and, if 'high stakes' are attached to the results in terms of the consequences of the publication of 'league tables', then we can expect teachers to teach to the test and restrict the curriculum accordingly.

Wood (1986) argued that educational measurement aims should be *"to devise tests which look at the individual and find out how well rather than how many"*. Wood's definition of educational measurement was that it:

- deals with the individual's achievement relative to himself rather than to others;
- seeks to test for competence rather than for intelligence;
- takes place in relatively uncontrolled conditions and so does not produce 'well-behaved' data;
- looks for 'best' rather than 'typical' performances;
- is not effective unless rules and regulations characteristic of standardised testing are maintained;
- embodies a constructive outlook on assessment where the aim is to help rather than condemn the individual.

There is also a need for distinction between competence and performance. Gipps (1994) said that *"Competence refers to what a person can do under ideal circumstances, while performance refers to what is actually done under existing circumstances. Competence includes the ability to access and utilise knowledge structures, as well as motivational, affective and cognitive factors that influence the response"*. Thus, according to Messick (1984), *"a student's competence might not be revealed in either classroom performance or test performance because of personal or circumstantial factors that affect behaviour"*.

It is important for educators to think of the impact on motivation and self-esteem if they use the wrong tools to assess their students. It is also important to find the assessment forms that are appropriate to individuals and to elicit the best performance from them. In order to do that, educators should be aware of the learning theories, which seek to understand why the students so often face difficulties and to align assessment with these theories. Furthermore, different types of assessment seem to encourage deep or surface approaches to learning (Struyven et al., 2002). For example, fixed response questions may encourage students to think dualistically even if designed to go beyond recall issues because, at the end of the day, students are asked to select one right answer. Therefore, it is argued that the content and style of a test have an important message to students about the nature of science and their intellectual development (Boud, 1995).

Understanding the psychological processes, which underpin learning, may provide useful information to avoid constructing questions which may be beyond any reasonable expectation of student's abilities. Crisp and Sweiry (2003) emphasised the importance of how a question is understood by subtle changes of certain aspects of a question such as diagrams or images, which are particularly salient and hence can come to dominate the mental representation that is formed. Many researchers (e.g. Oakhill, 1988; Davey, 1990) use the information processing model to explain the difficulties that pupils face when they try to answer negative questions.

The study of Lu and Suen (1995) showed that pupils' performance in different formats of assessment are related to their cognitive style. Pollitt et al. (2000) also addressed the problems related to the language barrier that students face when they study in a language which is not their mother tongue. They concluded that the problems are linguistic, contextual and cultural. However, language problems may simply reflect information overload (Selepeng and Johnstone, 2001).

Measuring instruments of the study

The following measuring instruments were employed to gather information from the pupils:

- Two cognitive tests:
  - Field Dependent/ Field Independent test
  - Convergent/ Divergent test
- Five chemistry paper-and-pencil tests

Measurement of field dependency

A version of the Witkin et al. (1971) group embedded figures test was used to determine an individual's degree of field dependency. It is almost identical to that used by Witkin and was used by Bahar (1999) in his study of cognitive structure. Known as the Hidden Figure Test (H.F.T.), it comprises twenty complex figures plus two additional, introductory figures that were used as examples. Simple geometric target shapes are presented on the back of a booklet. Pupils are required to recognise and identify one of the target shapes embedded within each of the complex figures. They do this by tracing its outline with a pen or a pencil. The main scoring scheme for the tests is to give one point for a correct simple shape embedded in a complex figure. The overall sum of the scores is the total mark that a student can gain. Thus the possible maximum score that can be obtained is 20.

Measurement of convergency/divergency

The Convergent /Divergent test consisted of 6 mini tests, described below.

- Test 1 was designed to find out the subjects' ability to generate words of the same or similar meaning to those given. At the beginning of the test an example was provided to show what the pupil was required to do. For example, if the word 'short' was given, a set of words such as 'abbreviated, limited, brief, concise, momentary, little, abrupt, petite, crisp, and compact' might be expected. This test included three questions and the time given for this test was 4 minutes.
- Test 2 asked the pupils to construct as many sentences as possible using four given specific words in each sentence, the words to be used in the form as given. Any sentences which did not make sense, received no credit. An example was provided at the beginning of the test and the time given for the test was 4 minutes.

- Test 3 is the only test which is not verbal. This is because there are some pupils who are pictorial learners and thinkers and, therefore, they perceive ideas more easily by pictures and diagrams. Thus, a pictorial test was included to give an opportunity to this type of student. In this test the student was required to draw up to five different pictures to relate to the idea of the given word. An example was given at the beginning of the test and 5 minutes was allowed.
- The purpose of test 4 was to see how many things the students could think of that are alike in some way. They were asked to write all the things that are round, or that are round more often than any other shape. An example was given at the beginning and 2 minutes was allowed for it.
- The objective of test 5 was to measure the student's ability to think of as many words as they could that begin with one letter and end with another. For example, students were asked about the words, which begin with the letter G, and end with the letter N. Names of people or places were not allowed and the time limit was 2 minutes.
- Test 6 aimed to find how many ideas the students could think of about a given topic. They had to list all the ideas they could about a topic whether or not it seemed important to them. An example was given at the beginning of the test and 3 minutes were allowed to complete this test.

The total time allocated for these six mini tests was 20 minutes. The researcher controlled the time limit for each test during the session. The test was translated (free translation) into Greek and two Language teachers checked the clarity of Greek carefully. The aim was to detect possible ambiguities and sources of confusion. In order to measure pupils' performance, one mark was given for every single correct response (Hudson, 1966). The test had been widely used in measuring extent of divergency by Bahar (1999).

Chemistry tests

Five paper-and-pencil chemistry tests were given to the pupils. Each chemistry test was designed to assess pupils by a range of question formats which tested the same knowledge and understanding in the same topics. The range of question formats that have been used in the project is shown in Table 1. This choice was made on the basis of the expectation that multiple choice questions would favour the convergent pupil, open-ended questions would be more congenial for the divergent pupils and that structural communication grids would appeal to both convergent and divergent pupils. For readers who are less acquainted with structural communication grids, examples of this format are shown in the appendix.

Because the teachers had to replace their classroom tests with the researcher's tests, the experimental tests were designed with the teachers' advice in mind, and an attempt was made to keep them short and appropriately demanding. The tests were constructed to be similar to the study questions of the Greek chemistry textbook (Liodakis et al., 1999), the Scottish Standard Grade Chemistry book (Renfrew and Conquest, 1995), and the book by Moore et al. (1998).

The tests were based on:

- Test 1: Atomic structure, classification of matter, solubility.
- Test 2: The periodic table and chemical bonds.
- Test 3: Mole concept.
- Test 4: Acids, alkalis, pH, neutralisation.
- Test 5: Solutions.

Table 1. Combination of different format questions.

	Test 1	Test 2	Test 3	Test 4	Test 5
Question Format	MC	SA	SA	MC	MC
	SA	SCG	SCG	SCG	SCG
					SA

MC: Multiple-Choice
SA: Short-Answer (Open-Ended)
SCG: Structural Communication Grid

Sampling method and administration procedures

The study was conducted in Greece during the school year September 2002 - May 2003 and 12 public upper secondary schools (*lykeio*) participated. There was more than one class in some schools and, therefore, the total number of classes was 23 and the total number of teachers was 12 (one teacher in each school). The classes were of different size: the smallest had 11 pupils and the largest had 29 pupils. Table 2 outlines the whole plan of the study.

Table 2. Schools and classes involved in the study.

Schools	Number of classes in each school	Number of pupils in each class				Total number of pupils
1	1	19				19
2	3	21	19	14		54
3	1	21				21
4	1	14				14
5	2	18	18			36
6	4	18	18	17	11	64
7	3	27	24	23		74
8	2	29	27			56
9	1	25				25
10	2	21	22			43
11	1	26				26
12	2	25	19			44
Total	**23**					**476**

It was decided to work with the pupils of Grade 10 (age 15-16) because, at that stage, pupils do not participate in national exams and teachers are more willing to be involved in research. Another very important reason to work with Grade-10 pupils was that all pupils have to attend chemistry lessons as, at that stage, pupils have not yet been split into arts or science streams. Thus, classes were heterogeneous, with pupils of different abilities and subject orientation.

The schools were not chosen at random because of the nature of the research. The researcher contacted teachers of different schools, before the beginning of the school year, and explained to them the plan of the study. The schools were selected in different geographical areas and of a different socio-economic background as much as possible. After receiving the teachers' agreement on the project, the researcher applied to the Greek Pedagogic Institute and Greek Ministry of Education for permission to have access to schools in order to administer the tests.

The schools and the number of pupils who participated in each test are summarised in Table 3. As can be seen from this table, the participating schools opted out of many tests, and this was unfortunate. The reasons for that were thought to be the lack of provision of organised training and educational studies for teachers as well as the very small amount of teaching time (just two 45 minute periods per week through the year). This makes teachers concerned about finishing the teaching units, and they are not willing to spend time on the evaluation of their teaching.

Table 3. Number of schools and pupils who participated in each test.

		Test 1	Test 2	Test 3	Test 4	Test 5
Schools	1	*				
	2			*		
	3	*				
	4	*				
	5	*				
	6		*	*		
	7	*				
	8	*	*		*	
	9	*	*		*	
	10	*	*			
	11					*
	12			*		*
Number of pupils		288	185	146	75	64

The cognitive tests were administered by the researcher. The class teachers in the various schools administered the chemistry tests.

Statistical methods used in the research

Validation and reliability of the instruments

The cognitive tests were based on well-established techniques. The field dependency test was almost identical to that of Witkin et al. (1971) test, while Bahar (1999) had developed and tested the materials used for the convergent/divergent test, based on standard tests for convergence-divergence.

Most statistical tests of reliability (other than test and re-test) indicate internal consistency. This procedure was not used in any of the chemistry tests because the tests consisted of sections having heterogeneous items assessing a mix of modes, degrees of difficulty and different depths of understanding. The chemistry tests were discussed with experienced class teachers in Greece to check face validity and minor adjustments were made.

Correlations

Both Pearson coefficient and Spearman's rho correlation between the different formats of the questions were calculated and were found to give similar values. However, because the distributions were frequently observed to deviate from a normal distribution, the Spearman's rho coefficient was more appropriate and this was used in all subsequent discussion.

Findings of the study

The H.F.T. results

The mean score of the H.F.T test was 7.8 (minimum = 0, maximum = 20) and the standard deviation was 4.2.

The CV/DV test results

The mean score of the convergent/divergent test was 47 (minimum = 0, maximum = 75) and the standard deviation was 10.7.

Chemistry test results

Table 4 shows the mean scores (%) and the standard deviations of each section of all the tests.

Table 4. Mean scores and s.d. of each format of each test.

	Test 1 (N=288)		Test 2 (N=185)		Test 3 (N=146)		Test 4 (N=75)		Test 5 (N=64)		
	MC	SA	SA	SCG	SA	SCG	MC	SCG	MC	SCG	SA
Mean	64.3	53.5	52.2	36.7	60.9	67.7	53.0	68. 0	67.5	68.6	50.8
s.d.	20.4	25.6	30.7	25.4	37.3	36.7	22.7	24.1	28.6	26.4	35.6

SA Short answer test
MC Multiple choice test
SCG Structural communication grid test

In general the Short Answer sections of each test were found to be more difficult than Multiple Choice and Structural Communication Grid sections of the tests. [Test 4 was a new test and was not part of the test battery used in our previous paper (Danili and Reid, 2005).]

Correlations between field dependent/independent cognitive scores and chemistry scores

Table 5 summarises the correlations for the field dependent/independent characteristic. It shows that the field dependent/independent characteristic correlated with pupils' performance in all the tests and in all the formats of assessment (although not always significantly). Being field independent seems to be a very important factor which influences whether pupils perform well in almost all types of assessments, and irrespective of the content of the questions. This result is consistent with the majority of the research in this field (e.g. Tinajero and Paramo, 1998; Danili and Reid, 2004; Tsaparlis, 2005) and the correlation values obtained here are very typical of previous work. Although significant, correlation values at this level indicate that only a small percentage of the variance is related to the field dependency skill. The short answer format of assessment favours field independent pupils more than the grid format of assessment. This is seen in tests 2 and 3, although the effect is not observed in test 5 with its smaller sample.

Table 5. Spearman's rho correlations for the field dependent/independent characteristic.

	Test 1 (N=288)		Test 2 (N=185)		Test 3 (N=146)		Test 4 (N=75)		Test 5 (N=64)		
	MC	SA	SA	SCG	SA	SCG	MC	SCG	MC	SCG	SA
rho	0.25	0.29	0.31	0.12	0.32	0.19	0.12	0.31	0.26	0.39	0.40
sig.	**	**	**	NS	**	*	NS	**	*	**	**

** Correlation is significant at the 0.01 level.
* Correlation is significant at the 0.05 level.
NS not significant

Although the field dependent/independent characteristic may develop naturally with experience, it may be difficult to teach someone to be field independent. However, attention should be given in the construction of the assessment to avoid confusion for those who are not able to separate the important information from the unimportant although, in some cases, the ability to separate the 'message' from the 'noise' may be an important skill to test. Thus, shredding (a process where a group of 'experts' scrutinise questions to ensure content validity, and to remove ambiguity and other errors) is a necessary process for quality assessment. Superficial clues, negative and double negative expressions, or subtle aspects, which can come to dominate the mental representations, should be avoided (Crisp and Sweiry, 2003; Johnstone, 2003).

Correlations between the convergent/divergent scores and chemistry scores

Table 6 summarises the correlations between the convergent/divergent characteristic scores and chemistry scores for different formats of assessment in the five chemistry tests. In general, the convergent/divergent characteristic correlated with pupils' performance in assessment, where language was an important factor to perform well (e.g. test 1, 2). Thus, in assessments that require pupils to have linguistic skills in order to elaborate and interpret a given text or to explain phenomena, ideas and concepts, or to describe differences, the convergent/divergent style is an important factor for pupils to perform well.

Table 6. Spearman's rho correlations for the convergence/divergence characteristic.

	Test 1 (N=288)		Test 2 (N=185)		Test 3 (N=146)		Test 4 (N=75)		Test 5 (N=64)		
	MC	SA	SA	SCG	SA	SCG	MC	SCG	MC	SCG	SA
rho	0.34	0.29	0.32	0.16	0.12	0.04	0.07	0.37	0.04	0.05	-0.13
sig.	**	**	**	*	NS	NS	NS	**	NS	NS	NS

** Correlation is significant at the 0.01 level.
* Correlation is significant at the 0.05 level.
NS not significant

It is possible to suggest that the short answer or open-ended questions favour divergent style pupils more than objective questions do. This is because, in short answer questions, pupils need to articulate their thoughts, and divergent pupils are better able to do this. In objective testing, if a question needs reading skill in order to elaborate and interpret a text given, then again the convergent/divergent style may be a very important factor for success. However, in algorithmic types of questions or in questions where there is more use of symbols and less use of words, such as test 3 and 5, or MC questions 1, 2, 3, 12 of test 1, the convergent/divergent characteristic does not relate to pupils' performance. In this case the format of the questions does not have an effect on pupils' performance.

From the above findings it seems that, in relation to the convergent/divergent characteristic, the chemistry content may well be a factor affecting the type of questions being asked, and may allow the question to be more easily tackled by, say, a divergent pupil. However, in almost all the tests the divergent pupils outperformed convergent pupils and, when there were short answer questions or open-ended questions, the differences in the performance between the convergent and divergent groups became larger. Hudson (1966) noted that the convergers tended to chose the sciences, but the divergers who did choose the sciences performed very well. The results here confirm the advantage of being divergent in examinations.

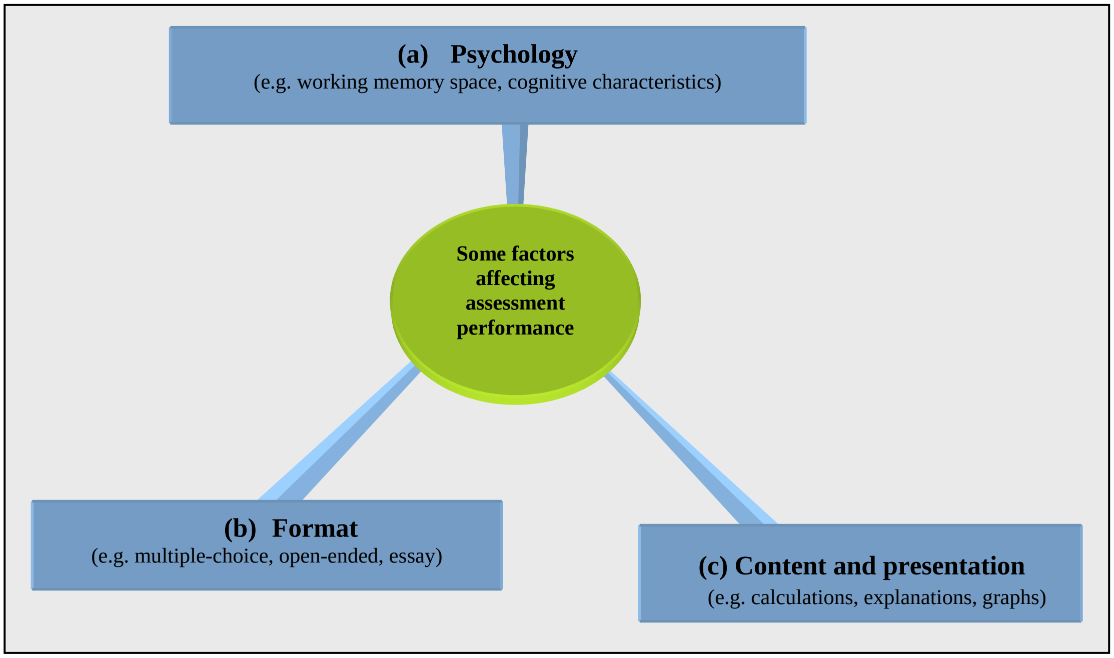
These outcomes are consistent with the work of Runco (1986), who indicated that there were particular domains of performance, for example art and writing, which were more strongly related to divergent thinking than other areas such as music and science. These results also might explain what Hudson (1966) pointed out: "*the convergence/ divergence dimension is a measure of bias, not a level of ability*". Thus, there is a tendency for those who are more divergent to feel more comfortable with some arts-orientated subjects. The closed nature of much early science teaching and learning may tend to attract those who are more convergent, while arts subjects offer opportunities for more extended thought, and attract those with good linguistic skills. However, if pupils who are good in linguistic skills choose science, it seems that they perform better than those who do not have such skills, because of their superiority in language. Linguistic skills such as comprehension and interpreting a scientific text are of paramount importance for reasoning in science (Byrne, et al., 1994). The results of Johnstone and Al-Naeme (1991) and of Field and Poole (1970) research offer some support for the claims.

In general, it seems that there is a relationship between the convergent/divergent characteristic and language. Although there is a debate as to whether thinking ability encourages divergence or divergent traits encourage thinking, it seems that the quality of a child's preschool language environment emerges as vital and, as Wittgenstein (1961) argued, the limits of one's language are the limits of one's world (Sutherland, 1992). Here is the importance of the teacher's role. The teacher should extend and challenge the child to go beyond where he would otherwise have been (Vygotsky, 1986). There is a need for teachers to encourage pupils to make their meaning explicit, and the use of the open-ended reports or essay assessments are useful tools for this.

Discussion and implications for good assessment practice

Based on all the above outcomes, some factors seem to have the potential to affect pupils' performance. Figure 1 reflects the observations from this study. There are, of course, other possible factors outside the scope of this work [see, for example, Taber (2003) for factors such as context, motive and purpose]. It also has to be noted that factors may well inter-relate with each other. For example, the test of convergence-divergence in this study correlated significantly, although at a low level, with the field dependent-independent test: those who are more divergent tend also to be those who are more field independent ($r = 0.19$, $p < 0.01$). This is not pursued further here.

Figure 1. Factors affecting psychological performance.



The study has raised many issues:

1. How do we decide about the validity of one format?
2. Is one format valid for one pupil but less so for another pupil?
3. Is there any format of assessment which is capable of being a more valid measure for most pupils?
4. What are we testing? Are we testing cognition or understanding of a particular discipline?
5. Do particular formats of assessment deskill the pupils?
6. Do particular formats of assessment frustrate pupils and therefore make them drop out of a subject or even out of study?

Assessment is a complex process. As Broadfoot and Black (2004) suggested,*"Educational assessment must be understood as a social practice, an art as much as a science, a humanistic project with all the challenges this implies and with all the potential scope for both good and ill in the business of education"*. In this situation Race (2003) suggests that *"Probably the best way to do our students justice is to use as wide as possible a mixture of assessment methods, … allowing students a range of processes through which to demonstrate their respective strengths and weaknesses"*. The issues are deeper as Thyne (1974), cited in Sanderson (1998), points out *"…it is axiomatic in the word of assessment that assessment tasks cannot measure 'cognition'…and the examiner must specify, at the outset, the performances to be accepted as evidence of Comprehending, or of Analysing, or whatever 'process' he wishes to assess, because examinations can measure only performance, not mental process"* Therefore, there are ethical issues about what formats of assessment need to be used to reflect pupils learning properly and, at the same time, to ensure a beneficial impact on teaching and learning practice (Gipps, 1994). It is important to be aware that testing may inadvertently favour a particular set of personal characteristics in the learner and thus test results may reflect possession of such characteristics as well as ability in the subject.

It follows that it is difficult to answer all the above questions fully, but it is possible to present the following guidelines for a good assessment practice:

1. Different formats may test different skills. Therefore we have to decide what we want to test. Do we want to test cognitive characteristics or to test knowledge and understanding?
2. It is impossible to design an assessment to suit each individual, and it is not wise to conduct all assessment by one method (e.g. objective testing, or open-ended).
3. The aims of the course must tie very closely to all aspects of assessment. If the aim of the course is the recognition of knowledge, then the test (probably multiple choice) must reflect that. If the aim is to transfer and apply knowledge, then problem solving, open-ended questions should be used. If the aim is to equip pupils with manual skills, then hands-on, or performance-based assessment should apply.
4. Teachers should be made aware that assessment is a skilful and demanding process, which has to be acquired through professional training. It is not just an appendage to teaching and learning.

Using a battery of different formats of assessment may help to ameliorate some of the problems outlined below:

- The use of objective tests may be of benefit to those pupils who, from nature or nurture, have not developed the cognitive processes needed to bring ideas together and present them clearly in open-ended formats. Also in a multicultural world, pupils may be assessed in a language which is not their mother tongue, and this means that may not perform well in some types of assessment where facility and subtlety of language is a limiting factor in their ability to produce an adequate answer.
- The use of open ended or problem-solving tasks helps more intellectually developed pupils to demonstrate their knowledge, their learning strategies, and to show their independence of thought. It is likely that objective tests constrain the more intellectually developed pupils and deprive them of opportunities to expand ideas.
- The use of oral examinations, open-ended assessments, essays, performance-based assessments, reports, portfolios and general alternative assessments encourages pupils to make their meaning explicit, to expand and enrich their vocabulary and their linguistic skills. Objective testing may deskill them linguistically.
- Assessment should not be punitive and judgemental but empowering and humane, especially at the school level when the pupils are forming their personality, building their self-esteem, and testing themselves in a different environment from their home. Assessment practice should support human needs rather than frustrate them. This means that assessment should encourage less successful pupils in their self-esteem, and help them to be less anxious about their performance, and therefore make them feel more comfortable in the school environment and stay longer in the school. After all we are human beings and we are entitled to make mistakes and to learn from them. Assessment for learning is an important issue (see Sorenson, 2000).

References

Al-Naeme F.F.A., (1988), *A study of some psychological factors affecting performance in chemistry at secondary and tertiary levels*, M.Sc. Thesis, University of Glasgow.

Al-Naeme F. and Johnstone A., (1991), Room for scientific thought?, *International Journal of Science Education*, **13**, 187-192.

Andersson B.R., (1990), Pupils' conceptions of matter and its transformation on classroom activities (age 12-16), *Studies in Science Education*, **18**, 53-85.

Atkins M., Beattie J. and Dockrell B., (1992), *Assessment issues in higher education*, Newcastle School of Education, University of Newcastle.

Bahar M., (1999). *Investigation of Biology student's cognitive structure through word association tests, mind maps and structural communication grids*, Ph.D. Thesis, University of Glasgow.

Bahar M. and Hansell M., (2000), The relationship between some psychological factors and their effects on the performance of grid questions and word association tests, *Educational Psychology: An International Journal of Experimental Educational Psychology*, **20**, 349-363.

Bennett S.N., (1973), Divergent thinking abilities–a validation study, *The British Journal of Educational Psychology*, **43**, 1-7.

Biggs J., (1996), Enhancing teaching through constructive alignment, *Higher Education*, **32**, 347-346.

Bodner G.M., (1991), I have found you an argument, *Journal of Chemical Education*, **68**, 385-388.

Boud D., (1995), *Assessment and learning: contradictory or complementary*? In P. Knight (Ed.), Assessment for Learning in Higher Education, London, Kogan Page Ltd.

Boud D., (2004), *Aligning assessment with long term learning needs*, Paper presented at the Enhancing the student experience in Scottish higher education. Constructive alignment of learning outcomes to assessment methods, Bridge of Allan, University of Stirling.

Broadfoot P. and Black P., (2004), Redefining assessment?, The first ten years of Assessment in Education. *Assessment in Education*, **11**, 7-26.

Byrne, M., Johnstone, A., and Pope, A., (1994), Reasoning in science: a language problem revealed? School *Science Review*, **75** (272), 103-107.

Child D. and Smithers A., (1973), An attempted validation of the Joyce-Hudson scale of convergence and divergence, *British Journal of Educational Psychology*, **43**, 57-61.

Crisp V. and Sweiry E., (2003, September), Paper presented at the British Research Association Annual Conference, www.ucles-red.cam.ac.uk, Edinburgh.

Danili, E. and Reid N., (2004), Some strategies to improve performance in school chemistry, based on two cognitive factors, *Research in Science & Technological Education*, **22**, 201-223.

Danili E. and Reid N., (2005), Assessment formats: do they make a difference? *Chemistry Education Research and Practice*, **6**, 204-212.

Davey B., (1990), Field dependence-independence and reading comprehension questions: task and reader interactions, *Contemporary of Experimental Education*, **52**, 119-206.

Entwistle A. and Entwistle N., (1992), Experiences of understanding in revising for degree examinations, *Learning and Instruction*, **2**, 1-22.

Fatt, J.P.T., (2000), Understanding the learning styles of students: implications for educators, *International Journal of Sociology and Social Policy*, **20**, 31-45.

Field T.W. and Poole M.E., (1970), Intellectual style and achievement of arts and science undergraduates, *British Journal of Educational Psychology*, **40**, 338-341.

Fryer M., (1996), *Creative teaching and learning*, London, Paul Chapman.

Gabel D., (1999), Improving teaching and learning through chemistry education research: a look to the future, *Journal of Chemical Education*, **76**, 548-553.

Gipps C., (1994), *Beyond Testing: towards a theory of educational assessment*, London, Routledge Falmer.

Hudson L., (1966), *Contrary Imaginations: a psychological study of the English schoolboy*, Great Britain, Penguin books Ltd.

Hudson L., (1968), *Frames of mind*, London, Methuen.
Johnstone A.H., (2003), *LTSN Physical sciences practice guide: effective practice in objective assessment.* Hull, LTSN.
Johnstone A.H. and Al-Naeme F.F., (1995), Filling a curriculum gap in chemistry, *International Journal of Science Education*, **17**, 219-232.
Johnstone A.H. and El-Banna H., (1986), Capacities, demands and processes: a predictive model for science education, *Education in Chemistry*, **23**, 80-84.
Kogan N., (1976), *Cognitive styles in infancy and early childhood*, New Jersey, Lawrence Erlbaum Associates, Inc.
Liodakis S., Gakis D., Theodoropoulos D., Theodoropoulos P. and Kallis A., (1999), *Chemistry A*' Lyceum Athens, Didaktika Biblion.
Lu C. and Suen H.K., (1995), Assessment approaches and cognitive styles, *Journal of Educational Measurement*, **32**, 1-17.
Messick, S., (1984), The psychology of educational measurement, *Journal of Educational Measurement*, **21**, 215-238.
Messick S., (1993), *The matter of style: manifestations of personality in cognition, learning, and teaching*, Princeton, NJ, Educational Testing Service.
Messick S., (1994), The matter of style: manifestations of personality in cognition, learning, and teaching. *Educational Psychologist*, **29**, 121-136.
Moore J., Stanitski C., Wood J., Kotz J. and Joesten M., (1998), *The chemical world concepts and application*s (2nd ed.), New York, Saunders College.
Nurrenbern S. and Pickering M., (1987), Concept learning versus problem solving: is there a difference? *Journal of Chemical Education*, **64**, 508-510.
Nuttall D.L., (1972), Convergent and divergent thinking, *The British Journal of Educational Psychology*, **43**, 1-7.
Oakhill J., (1988), *The development of children's reasoning ability: information processing approaches*, in K. Richardson, and Sheldon S. (Ed.), Cognitive development to adolescence, LEA/OU.
Orton A., (1992), *Learning mathematics issues, theory and classroom practice*, London, Cassell Education Series.
Osborne M., (2004) *Constructive alignment of assessment methods with learning outcomes*, An overview for the assessment workshop on 27th February 2004. Paper presented at the Enhancing the student experience in Scottish higher education, Constructive alignment of learning outcomes to assessment methods, Bridge of Allan.
Osborne R.J. and Cosgrove M.M., (1983), Children's conceptions of the changes of state of water, *Journal of Research in Science Teaching*, **20**, 825-838.
Pollitt A., Marriott C. and Ahmed A., (2000), *Language, contextual and cultural constraints on examination performance*, Paper presented at the International Association for Educational Assessment, Jerusalem, Israel, http://www.ucles-red.cam.ac.uk/conferencepapers.htm.
Race P., (2003), *Designing assessment to improve Physical Sciences learning – exams*, http://dbweb.liv.ac.uk/ltsnpsc/guides/designing_assessment/exams.htm [2003, 6/10/2003].
Renfrew R. and Conquest N., (1995), *Standard Grade Chemistry*, London, Hodder and Stoughton.
Riding R.J. and Cheema I., (1991), Cognitive styles: an overview and integration, *Educational Psychology*, **11**, 193-215.
Runco M.A., (1986), Divergent thinking and creative performance in gifted and non gifted children, *Educational and Psychological Measurement*, **46**, 375-383.
Sanderson P., (1998), *Cognitive process and discourse in the assessment of written text*, Paper presented at the European Conference for Educational Research, University of Ljubljana, Slovenia, http://www.ucles-red.cam.ac.uk/conferencepapers.htm.
Sawrey B.A., (1990), Concept learning problem solving: revisited, *Journal of Chemical Education*, **67**, 253-254.

Selepeng D. and Johnstone A.H., (2001), A language problem revisited, *Chemistry Education Research and Practice*, **2**, 19-29

Shepard L., (1992), *What policy makers who mandate tests should know about the new psychology of intellectual ability and learning*, In B. Gifford, and O'Connor, M. (Ed.), Changing Assessments: alternative views of aptitude, achievement and instruction, London, Kluwer Academic Publishers.

Sorenson P., (2000), *Drowning in numbers? The need for formative assessment*, in Sears, J and Sorenson, P, Issues in Science teaching, London, Routledge Falmer, pp. 123-132

Struyven K., Dochy F. and Janssens S. (2002), *Students' perceptions about assessment in higher education: a review*, Education-line, http://www.leeds.ac.uk/educol/documents/000025.htm.

Sutherland P., (1992), *Cognitive development today: Piaget and his critics*, London, Paul Chapman Publishing Ltd.

Taber K.S., (2003), Examining structure and context – questioning the nature and purpose of summative assessment, *School Science Review*, **85**, 35-41.

Terrell S.T., (2002), *The use of cognitive style as a predictor of membership in middle and High School Programs for the academically gifted*, Paper presented at the Annual meeting of the American Educational Research Association, New Orleans, Louisiana.

Thyne R.L., (1974), *Principles of examining*, London, University of London Press.

Tinajero C. and Paramo F.M., (1998), Field dependence-independence cognitive style and academic achievement: a review of research and theory, *European Journal of Psychology of Education*, **13**, 227-251.

Tsaparlis G., (2005), Non-algorithmic quantitative problem solving in university physical chemistry: a correlation study of the role of selective cognitive factors, *Research in Science and Technological Education*, **23**, 125-148.

Vygotsky L., (1986), *Thought and language*, (A. Kozulin, trans.), Cambridge, MA, The MIT Press (Originally published in Russian in 1934).

Witkin H.A. and Asch S.E., (1948a), Studies in space orientation III. Perception of the upright in the absence of visual field, *Journal of Experimental Psychology*, **38**, 603-614.

Witkin H.A. and Asch S.E., (1948b), Studies in space orientation IV. Further experiments on perception of the upright with displaced visual fields, *Journal of Experimental Psychology*, **38**, 762-782

Witkin H.A., Dyk R.B., Paterson H.F., Goodenough D.R. and Karp S.A., (1974), *Psychological differentiation - studies of development*, New York, Wiley.

Witkin,H.A. and Goodenough D.R., (1981), *Cognitive styles: essence and origins*, New York, International University Press.

Witkin H.A., Goodenough D.R., Moore C.A. and Cox P.W., (1977), Field-dependent and field-independent cognitive styles and their educational implications, *Review of Educational Research*, **47**, 1-64.

Witkin H.A., Oltman P.K., Raskin E. and Karp S.A., (1971), *Embedded figures test, children's embedded figures test, group embedded figures test: Manual*, Palo Alto, CA, Consulting Psychologists Press.

Wittgenstein L., (1961), *Tractatus logico-philosophicus*, London, Routledge and Kegan Paul.

Wood R., (1986), *The agenda for educational measurement*, In D. Nuttall (Ed.), Assessing educational achievement, London, Falmer Press.

Appendix:

Structural Communication Grids

This is an objective form of assessment, which has distinct advantages over conventional multiple choice testing. The grid is a database on which a number of questions are based. The student is not told how many pieces of data are required to answer each question and so every piece has to be read and weighed up for its relevance to the question. The element of guessing is greatly reduced and students are seeking classification patterns rather than recognising one correct answer out of four or five. Successive questions will need some of the grid elements already used, and so answers cannot be correctly found merely by elimination. This process may involve subtlety of language interpretation, which may appeal more to the field independent and divergent students.

Each question is answered by an array of letters corresponding to the letters in the corner of each box in the grid, which has been selected by the pupil. Two examples are given below.

1. Each box in the grid below refers to a compound. Look at the boxes and answer the questions that follow. *(Boxes may be used as many times as you wish)*

A Sodium oxide Na_2O	*B* Lead nitrate $Pb(NO_3)_2$	*C* Phosphorus trioxide P_2O_3
D Barium iodide BaI_2	*E* Calcium oxide CaO	*F* Sodium nitrate $NaNO_3$
G Sulphur dioxide SO_2	*H* Magnesium sulphate $MgSO_4$	*I* Nitrogen dioxide NO_2

Select the box(es) that contain compounds which:

a) Produce alkaline solutions when dissolved in water..............................
b) Produce acidic solutions when dissolved in water ....
c) Cause acid rain………………………
d) Can react with the salt in box D and give a precipitation reaction……………..

2. Each box in the grid below refers to an element. Look at the boxes and answer the questions that follow. *(Boxes may be used as many times as you wish)*

A The element with the electron arrangement: 2.8.3	*B* Sodium	*C* Ar
D Magnesium	*E* The element that is a brown liquid at room temperature	*F* The element that has one electron in each atom
G The element with atomic number 19	*H* Chlorine	*I* Nitrogen

Select the box(es) that contain:

a) Elements that are gases in room temperature………………
b) Elements in the same group of the periodic table………………
c) Two elements that will combine to form an ionic compound with the formula X_3Y_2 ....
d) Elements that form a covalent compound with the element which is in box F ………………