The Development and Validation of a Categorization of Sources of Misconceptions in Chemistry

Abstract: The purpose of this research is to devise a classification of sources of nonexperiential misconceptions, characteristic of learning Chemistry, and to establish experimental data verifying that categorization. The categories of error sources were developed based on a review of the literature, from the 23 years of teaching experience of the investigator, and from discussions with other experienced Chemistry teachers. A summary of the background research is presented below. Of necessity, it is considerably abbreviated, but the majority of the studies used in the categorization have at least been referenced in order to provide the reader with the resources used.

A misconception is defined as a mental representation of a concept which does not correspond to currently held scientific theory. The term misconception, as used in this study, refers exclusively to those misconceptions concerning science. Misconceptions have increasingly been the focus of research in Science Education over the past decade, with the vast majority of studies focusing on the discipline of Physics (Clement, 1983; McCloskey, 1983; Champagne, Gunstone, and Klopfer, 1985; Duit, 1987). These experiential misconceptions are also referred to as alternative conceptions (Driver, 1983), intuitive conceptions (Clement, 1983), or naive conceptions (McCloskey, 1983). In each case the author is describing a concept which has been understood, at least to some extent, through everyday experience and interaction with the phenomenon involved. Examples of experiential misconceptions occur in connection with phenomena such as motion, energy, and gravity. Misconceptions pertaining to chemical phenomena, however, are fundamentally different because the existence of atoms and molecules are not directly within the realm of everyday experience. Misconceptions pertaining to these more abstract phenomena result from some instructional experience, within or outside of the classroom, including self instruction. For purposes of this study, they will be called instructional misconceptions. Because the nature of misconceptions in Chemistry is basically different from experiential misconceptions in both origin and kind, investigation should prove fruitful in identifying their sources and providing a mechanism for addressing them directly.

Keywords: concept formation, educational methods, research methodology, concept formation, misconceptions, language processing, cognitive processes, error patterns, protocol analysis

General School Subject: chemistry
Specific School Subject: college prep. chemistry
Students: college bound

Macintosh File Name: Skelly - Chemistry
Release Date: 12-16-1993 C, 11-6-1994 I

Publisher: Misconceptions Trust
Publisher Location: Ithaca, NY
Volume Name: The Proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics
Publication Year: 1993
Conference Date: August 1-4, 1993
Contact Information (correct as of 12-23-2010):
Web: www.mlrg.org
Email: info@mlrg.org


Note Bene: This paper is part of a collection that pioneered the electronic distribution of conference proceedings. Academic livelihood depends upon each person extending integrity beyond self-interest. If you pass this paper on to a colleague, please make sure you pass it on intact. A great deal of effort has been invested in bringing you this proceedings, on the part of the many authors and conference organizers. The original publication of this proceedings was supported by a grant from the National Science Foundation, and the transformation of this collection into a modern format was supported by the Novak-Golton Fund, which is administered by the Department of Education at Cornell University. If you have found this collection to be of value in your work, consider supporting our ability to support you by purchasing a subscription to the collection or joining the Meaningful Learning Research Group.

----- -----
The Development and Validation of a Categorization of Sources of Misconceptions in Chemistry
Kathleen M. Skelly, University of Massachusetts at Lowell
and Dana Hall School, Wellesley, MA 02181 USA

OVERVIEW

The purpose of this research is to devise a classification of sources of nonexperiential misconceptions, characteristic of learning Chemistry, and to establish experimental data verifying that categorization. The categories of error sources were developed based on a review of the literature, from the 23 years of teaching experience of the investigator, and from discussions with other experienced Chemistry teachers. A summary of the background research is presented below. Of necessity, it is considerably abbreviated, but the majority of the studies used in the categorization have at least been referenced in order to provide the reader with the resources used.

A misconception is defined as a mental representation of a concept which does not correspond to currently held scientific theory. The term misconception, as used in this study, refers exclusively to those misconceptions concerning science. Misconceptions have increasingly been the focus of research in Science Education over the past decade, with the vast majority of studies focusing on the discipline of Physics (Clement, 1983; McCloskey, 1983; Champagne, Gunstone, and Klopfer, 1985; Duit, 1987). These experiential misconceptions are also referred to as alternative conceptions (Driver, 1983), intuitive conceptions (Clement, 1983), or naive conceptions (McCloskey, 1983). In each case the author is describing a concept which has been understood, at least to some extent, through everyday experience and interaction with the phenomenon involved. Examples of experiential misconceptions occur in connection with phenomena such as motion, energy, and gravity. Misconceptions pertaining to chemical phenomena, however, are fundamentally different because the existence of atoms and molecules are not directly within the realm of everyday experience. Misconceptions pertaining to these more abstract phenomena result from some instructional experience, within or outside of the classroom, including self instruction. For purposes of this study, they will be called instructional misconceptions. Because the nature of misconceptions in Chemistry is basically different from experiential misconceptions in both origin and kind, investigation should prove fruitful in identifying their sources and providing a mechanism for addressing them directly.

THEORETICAL FRAMEWORK FOR MISCONCEPTION RESEARCH:
The Constructivist View of Learning

The constructivist view of learning provides the theoretical umbrella for misconception research (Strike and Posner, 1985; Driver and Oldham, 1986; Pines and West, 1986; Novak,
Constructivism is an effort to integrate the psychology of human learning and the epistemology of knowledge production as articulated by modern philosophers of Science, notably Kuhn (1970) and Lakatos (1972). For the constructivist, knowledge is seen as the product of the knower’s individual construction of his/her subjective reality. Hence, an individual’s knowledge is not a fixed entity, but rather is under continual development and restructuring.

This view of knowledge changes the role of the teacher, from one who transmits already structured knowledge, into one who facilitates the learner’s construction of meaning and understanding (Wittrock, 1985). There are two basic issues in the constructivist view of learning: the prior knowledge of the learner, and the active nature of the learning process (Magoon, 1977; Wittrock, 1985; Driver and Oldham, 1986; Scott, 1987; Fosnot, 1989; and vonGlasersfeld, 1988, 1989a, 1989b). These issues are the key areas of research which have emerged in Cognitive Science, and continue to be articulated by work in this field.

Construction of new knowledge is a highly interactive process with the individual exchanging information with other people and with his/her environment. The strongest evidence that knowledge acquisition may be accomplished by a process of individual construction, as opposed to simple assimilation of concepts previously organized by teacher or text, is found by examining the knowledge structures acquired by learners. Some of the less conventional assessment techniques, such as qualitative verbal explanations or laboratory practical situations, can reveal highly individual understandings on the part of the learner, understandings which often do not correspond to currently held scientific theory (Driver and Oldham, 1986; Hein, 1991). Misconceptions presumably are not taught directly or intentionally, but they do sometimes result from one individual’s making meaning of a situation.

THE LITERATURE ON MISCONCEPTIONS: DEVELOPING THE CLASSIFICATION

The extensive study of misconceptions over the past decade has produced an large body of related literature. Pfundt and Duit (Duit, 1991, 1987) have compiled, and continuously update, a bibliography on all aspects of the subject. They report that initial research focused on theoretical aspects of the issue, and then on descriptive studies identifying particular misconceptions. Duit indicated that the vast majority of studies are in the area of Physics, with less than 10% of the research in Chemistry, and even fewer in Biology. In general, the literature does not make distinctions between different types of misconceptions. This is not surprising when one considers that the large majority of current research addresses only the experiential misconceptions, characteristically found in Physics. It is the position of this investigator, however, that there are various kinds and various sources of misconceptions,
notably experiential versus instructional. The examination of the available literature synthesizes and outlines patterns in the nature and sources of misconceptions. The classification which follows is a synthesis of the findings. The initial classification, based on source of misunderstanding, divides misconceptions into two principal categories, namely those which arise from life experiences (Experiential Misconceptions), and those which result from some instructional process (Instructional Misconceptions) (Pines and West, 1983; Perkins and Simmons, 1988).

**Experiential Misconceptions**

Misconceptions which derive from everyday experience generally emerge before instruction takes place. For the purpose of this study, this type of misconception will be referred to as experiential misconceptions. These experiential misconceptions are extremely resistant to change (Posner et al, 1982; Hashweh, 1986; Perkins and Simmons, 1988). It is important to note here that these misconceptions do have a basis in logic. They do help to explain the way the world works for practical purposes. They explain things as they seem to be in the real world. That they “make sense” no doubt contributes to their robustness and resistance to change. Hashweh (1986) proposed that these intuitive conceptions may be proceduralized, by which he meant that they are stored in the individual’s conceptual framework in nonverbal or subconscious form. Although this would, of course, be relatively difficult to substantiate, it does fit well with observable characteristics of intuitive conceptions and provides one explanation for their resistance to change.

Experiential misconceptions are observed frequently in domains within Physics, most conspicuously in mechanics, because everyday life provides people with endless opportunities to observe and interact with objects in motion, in a real and "friction-filled" world. Common misconceptions are surprisingly widely held, even by the adult population. In one disheartening study (Lawrenz, 1986), elementary school teachers demonstrated extensive intuitive misunderstanding of concepts involving mass, gases, and basic mechanics. Champagne, Klopfer, and Anderson (1980) identified a number of “common sense” beliefs among Physics students, which are qualitatively different from the constructs of Newtonian mechanics. For example, a common belief is that objects which are under a constant force will move at a constant velocity. Studies by McCloskey, Caramazza, and Green (1980), Caramazza, McCloskey and Green (1981), and McCloskey (1983) indicate that most freshman college students, including those who have finished one or more years of Physics, as well as those with no Physics background, do not have a qualitative understanding of the most fundamental principles of mechanics.
Further subdivision of this category is possible, but fairly subtle. Fisher and Lipson (1986), for example, have made distinctions between experiential misconceptions based on specific flaws in reasoning. But for purposes of this study, it is sufficient to cluster experiential misconceptions into one group.

**Instructional Misconceptions**

The principal focus of this study is the second category of misconceptions, which encompasses those arising from an instructional situation, either formal or informal. This classification has been further subdivided by searching for an answer to the question, "What goes wrong in the learning process?" The answer to this question is crucial to a new view of Science Education. The most prevalent answers to this question, from the studies which have been examined, include the following:

1. Language related misunderstanding and misinterpretation of vocabulary, analogies, symbols, or overall meaning.
2. A deficient knowledge base on the part of the learner.
3. Overtaxing the short term memory of the learner.
4. A mismatch of the cognitive demands of the subject matter with the cognitive level of development of the learner.
5. Choice of mental strategies inappropriate to the subject matter.
6. Low standards of epistemology on the part of the student.

While some overlap may be found, each of these is sufficiently distinct to be examined separately.

**1. Language Related Misconceptions**

Language is the greatest source of misunderstanding in learning. Central to the constructivist view of learning is the belief that meaning cannot be simply transferred intact by one person to another. In VonGlasersfeld’s (1988) "technical model" of communication, he explained that the physical signals which travel from one person to another do not actually contain an entire meaning. They are, rather, a set of instructions to select particular meanings from the receiver’s repertoire of meanings. So the interpretation must be constructed by the receiver, and is a function of his/her sets of meanings. The process of developing and tuning the meanings of words and linguistic expressions is a life-long chore for everyone (Ausubel, Novak, and Hanesian, 1978). It is very important to make the point that the use of language in teaching is far more complicated than has been believed. "Telling" is not enough, because
understanding is not a matter of passively receiving, but of actively building up meaning. An examination of some of the meanings which students construct makes this all too apparent.

**Vocabulary**

Vocabulary related misconceptions are one of the greatest problems in Science learning. Students are adept at learning definitions from books and giving answers in appropriate technical language which makes it appear that they have an accurate understanding, when in fact they do not (Novak, 1987). Carey (1986) pointed out that many units of instruction in Science introduce more new vocabulary words than typical units of instruction in a foreign language. To complicate this further, in learning the foreign language, the student presumably is just learning a new word for a concept s/he already understands. Whereas in Science, the student is learning an entire concept including the vocabulary to describe it.

One aspect of the vocabulary problem is the scientific use/meaning of words which have a very different, usually less specified, meaning in everyday language (Clough and Driver, 1986). In one study (Jacobs, 1989), a group of undergraduate introductory Physics students were tested for their understanding of 25 vocabulary words with a standard English meaning and a specialized Physics meaning. Results indicated that students expressed confidence that they knew the meanings of an average of 15 words (of the 25), for which they were not using the correct meaning at all. In another study (Ribeiro, Pereira, and Maskill, 1990), 14 fourth year undergraduate Chemistry students were interviewed regarding the spontaneity of several common chemical reactions. All but two of the students answered the question as to whether each reaction was spontaneous using the everyday meaning of the term. The students were then asked to calculate the value of the Gibbs Free Energy to make a prediction of spontaneity. Twelve of the 14 could use the formula \( \Delta G = \Delta H - T \Delta S \), but when confronted with the mathematical evidence that two of the reactions which they had predicted as nonspontaneous were indeed spontaneous, only 3 additional students changed to the correct prediction.

Also, students are generally not aware of how specific and complete the ancillary knowledge of a concept must be to completely define that concept, while teachers make the assumptions that students do know this. Consider the concept of acceleration. To be completely articulated, it must be specified a) in succinct mathematical form: \( a = \frac{\Delta v}{\Delta t} \), b) informally by a verbal description, c) by describing its most salient features, d) procedurally, by illustrating the ways in which it operates, and e) by specifying any independent variables (Reif, 1985). Teachers themselves rarely articulate a concept with sufficient elaboration to facilitate meaningful understanding. A study by Veiga, Pereira, and Maskill (1989) demonstrated clearly that student misconceptions regarding heat, temperature, and energy were embedded in the language used by the teachers to communicate the concepts. Thus, common misconceptions were
unintentionally being continually reinforced. The students' meanings for the words simply did not correspond to the teachers' meanings.

Particular examples of vocabulary related misunderstandings can be found in all disciplines of science. The examples chosen here illustrate problematic areas in the learning of Chemistry. The terms atom and molecule are profoundly misunderstood. Griffiths and Preston (1989) found that twelfth grade students' misconceptions included thinking that atoms are solid spheres, are changeable in size, and some even thought of them as living. Students' concepts of the size, shape, and composition of molecules were seriously distorted. Both the Griffiths and Preston (1989) study and one by Ben-Zvi, Eylon and Silberstein (1986) found that students transferred properties of macroscopic samples of substances to individual atoms or molecules of a substance. This included such properties as expansion and contraction with temperature changes, malleability, and changes of state. Andersson's (1986) findings on students' explanations of observations of macroscopic changes in chemical reactions to show such beliefs as transmutation of atoms, substances' ceasing to exist, and substances freely entering or leaving the reaction chamber.

Chemical equilibrium presents particularly unique opportunities for misconceptions in this category. Several studies have investigated misunderstandings of basic vocabulary in this topic (Johnstone, MacDonald and Webb, 1977; Gussarsky and Gorodetsky, 1988.) The topic is sufficiently complex to provide a rich source of information, and has been selected for this study for that reason.

**Analogy and Metaphor**

The use of analogy and metaphor can be a powerful learning device in science, but this can also lead to misconceptions. An analogy is an explicit comparison between the structures of two different domains: the source domain of the analogy, and the target domain to be explained (Duit, 1991). An analogy focuses on the identity of parts or structures which the two domains have in common, while often overlooking the differences between them. A particular analogy seems to work only for a very specific target, because the structure of the analogy must be a close match to the target concept (Gentner and Gentner, 1983; Dupin and Johsua 1989; Gilbert, 1989). In some cases, the analogies used were more complex than the phenomena they were intended to explain (Curtis and Reigeluth 1984). It is not surprising that they sometimes did not have a positive result. Although it is beyond the scope of this paper to elaborate in greater detail, several studies of interest include Rigney and Lutz (1976), Raven and Cole (1978), Gabel and Sherwood (1980), and Webb (1985). The importance and usefulness of analogies in the process of knowledge construction becomes apparent in the context of schema theory as articulated by Rumelhart (1980) and Rumelhart and Norman (1981), and will be
discussed further in the next section, on the category of misconceptions related to the prior knowledge of the learner.

The ways in which analogies and metaphors promote concept formation differ from each other. The metaphor is primarily a tool to activate imaginative thinking. In comparing two very different phenomena, the mind is forced to think in terms of new relationships. Metaphors can enable the learner to let go of tenaciously held theories, and see a familiar phenomenon from a completely different point of view. They seem to provide an important bridge between the cognitive and affective domains, forcing creative thinking while motivating and interesting students (Gowin, 1983; Howard, 1989; Duit, 1991). The problem arises when students take metaphors too literally, in which case they can interfere with new learning. When metaphors are used, teachers have to monitor the meaning conveyed carefully (Howard, 1989; Duit, 1991).

Symbols

Misinterpretation of symbolic knowledge can cause considerable confusion. Symbols, which may or may not relate to the name of the concept which they represent, are a constant problem to Chemistry students (and no doubt to other Science students as well.) To name only several example, "m" is used to represent both milli- and mass; "M" represents both molar mass and molarity; and "n" represents the number of moles, whereas "N" stands for the number of objects in a mole as well as normality, a term confusing enough in its own right. Symbols for the elements provide more pitfalls, with some derived from the English name for the element, such as S for sulfur, while others are taken from the old Latin name used by the alchemists: Na (from natrium) for sodium, or Hg (from hydragryrum) for mercury. In addition, there is the question as to exactly what the symbol represents: an atom, a molecule, a gram, or a mole. However, difficulties which new students encounter with these problems have not been studied to any great extent.

2. Misconceptions Related to Prior Knowledge

The learner's prior knowledge is the vehicle by which s/he processes new information. It is a most important variable in the success of learning science. If the learner's prior knowledge needed to process new information is incomplete, the knowledge gaps will result in confusion, inaccurate reasoning, and eventually in the formation of misconceptions. If the learner's prior knowledge structure contains misconceptions, these can cause further faulty reasoning and incorrect concept formation (Rumelhart, 1980; Rumelhart and Norman, 1981).

It is in the schema formation process that analogies can be a powerful tool, providing that the learner is familiar with the analog domain. But research shows that it cannot be
assumed that a previously taught concept may be used as an analogy, because students may hold major misconceptions related to the concept (Gabel and Sherwood, 1980; Gentner and Gentner, 1983; Clement, 1987).

Research in science education has provided considerable evidence for the effects of prior knowledge on concept acquisition. The role of factual knowledge has proved to be very important in the reasoning process, so much so that failure to demonstrate the higher level reasoning critical to correct concept acquisition in science may well be due to lack of knowledge rather than lack of formal reasoning capacity (Goodstein and Howe, 1978; Linn, 1980, 1982; Linn, Pulos, and Gans, 1981; Linn, Clement and Pulos, 1983).

Champagne, Klopf er, and Anderson (1980) designed a study to identify the preinstructional knowledge of first year college physics students in mechanics, mathematics, and reasoning skills, and to correlate these variables with the students’ success in learning classical mechanics. They found prior knowledge of mechanics to be a poor predictor of achievement in college physics only because almost all students demonstrated low prior knowledge, even those who had studied high school physics. However, mechanics achievement correlated with math skills at .55 (p < .001). They hypothesize that students with weak math skills have to focus more of their attention on the mathematical procedures, and correspondingly less on the mechanics. In fact, analysis of class observations showed that fifty percent of the questions asked in class involved the mathematics rather than the mechanics.

Prior knowledge in a domain promotes higher level reasoning and more sophisticated learning, and this finding holds for subjects of all ages (Chi, Glaser and Rees, 1982; Linn, 1986.) It also helps to explain why reasoning ability continues to develop and increase even when subjects have reached the Piagetian stage of formal reasoning: their knowledge base continues to broaden, deepen and become increasingly well organized (Linn, 1986.)

3. Misconceptions Related to Overtaxing the Short Term Memory of the Learner

Overtaxing of the short term memory (STM) capacity, the greatest limitation of the human brain, results in excessively high cognitive load. The result is that the learner starts reasoning from incomplete knowledge. Simon (1974) researched the size of a short term memory unit and the number of such units which can be held simultaneously in STM. He found that as familiarity with the units increased, it was possible to increase the size of a unit from a single word to a familiar phrase. The capacity of STM appears to range from five to seven units or "chunks." The net content of STM is increased as chunk size is increased. For the beginner, however, chunk size is small, and working memory space is severely limited. Any overload results in some critical information being missing in the reasoning process, resulting in faulty
learning or no learning at all. A teacher’s working memory is using knowledge which is already organized, and s/he attempts to transmit a fully organized set of ideas. But the learner has not yet created an organization for him or herself, and cannot receive the information intact. Problem solving tasks should require no more than seven, and probably no more than five “chunks” of information. And initially those chunks have to be quite small, increasing only as some mastery and automation has taken place.

Sweller (1988) points out another aspect of STM overload, that of too many mental processes competing for available processing capacity. Activities which require too high a load on working memory, when much of that load does not assist in the development of schema of deep understanding of the topic, has the effect of impairing concept formation. Therefore, construction of accurate schemas should proceed the use of these schemas in problem solving in order to avoid incorporating misconceptions in the learner’s knowledge structure. A similar phenomenon was observed by this writer in interviewing first year university nursing students. The students studied were using algorithms to solve equilibrium buffer problems before they had any understanding of the nature of buffers. Probing questions revealed that many students had confused the concept of pH with the concept of $K_A$ as a direct result of cognitive overload.

4. Mismatch of Cognitive Demands of Subject Matter with Developmental Level of Learner

The content of Chemistry and Physics courses, in particular, and the methods normally used in teaching these subjects, require that the student operate at Piaget’s “formal” level of reasoning if s/he is to comprehend the material (Herron, 1975; Sayre and Ball, 1975; Howe and Durr, 1982). Piagetian theory predicts that students should begin the transition to the formal operational stage of development around the age of twelve, and be fully formal operational by the age of fifteen. But a large body of research indicates that most adolescents and young adults do not appear to be at the formal operational level of development (Karplus and Karplus, 1970; Karplus and Peterson, 1970; Karplus and Karplus, 1972; Chiapetta, 1976; Lawson, 1985).

An important issue with formal reasoning is that it is difficult to measure accurately (Linn, 1982; Lawson, 1983). Successful performance on a formal reasoning test does imply formal reasoning capacity, but initial unsuccessful performance does not necessarily mean the converse. It has been shown repeatedly that, although an individual has the capacity for formal reasoning, s/he does not necessarily use it in a particular instance (Karplus, Karplus, and Wollman, 1974; Linn, 1982; Lawson, 1985). A student must have the needed prior knowledge to elicit formal reasoning.
A second issue, the more important for this particular study, is that the capacity for formal reasoning has been shown to be the best predictor of achievement in science tasks where advanced reasoning is required, for example learning requiring control of variables, proportional reasoning, or correlational reasoning (Karplus and Karplus, 1970; Karplus and Peterson, 1970; Champagne, Klopfer, and Anderson, 1980; Howe and Durr, 1982; Lawson, 1982, 1985). If formal reasoning is needed to understand certain material, and the student does not have that capacity, s/he will not do well, and in all probability will experience frustration and loss of confidence. The problem is a particular one for the Chemistry teacher, because the subject matter clearly requires formal reasoning, and a large percentage of the student population is in a transitional state between concrete and formal operations, and some significant percent has not even started to make the transition.

Further problems for the student of Chemistry are caused by the introduction of very abstract concepts at too early an age. Young students must do their best to make sense of them in their own way, but will later find it difficult to identify, much less correct, their resulting misconceptions when they start their formal study of Chemistry. The extensive misunderstanding of atoms and molecules which is observed in older students (Griffiths and Preston, 1989) probably started in the early grades. Driver (1983) and Osborne and Freyberg (1985) present a vast body of data illustrating the extent to which children misunderstand abstract concepts such as atoms. Lawrenz’s (1986) data indicates that their teachers have none too clear an understanding of those concepts themselves. When one examines Andersson’s (1986) study of student explanations of the abstract changes taking place in a chemical reaction on the molecular level, one might well ask why seventh and eighth graders are being asked such a question. The same could be asked about studies involving the mole concept with ninth graders (Chiappetta and McBride, 1980). When very abstract concepts are introduced quite early, at which time children have neither the formal reasoning capacity nor the advanced mathematical skills to process the information correctly, misconceptions will be formed. These misconceptions are almost certainly avoidable, but once formed they are difficult to change.

5. Misconceptions Due to Inappropriate Mental Strategies

A fifth cause of misconceptions is the learner’s choice of mental strategies inappropriate to the subject matter. By far and away the strategy most frequently resorted to by students is rote learning. Unless the learner develops a deep understanding of a concept before memorizing, the concept will be assimilated in superficial form without the critical analysis necessary to form any schema at all, much less an accurate one. The underlying cause is often the course requirements of covering too much material too quickly (Johnstone, 1984). The cause could also be excessive cognitive demands of too much problem solving activity before
concepts have had adequate construction (Glaser, 1984; Sweller, 1988). Another cause could be the common practice of teaching problem solving through algorithms (Bodner, 1987). Such algorithms are most appropriate for experts, but they should not be taught to beginners directly. Rather the students should develop such techniques themselves as a result of structuring their own knowledge (Frank, Baker, and Herron, 1987).

6. Misconceptions Due to Students’ Standards of Epistemology

The final category mentioned here is that of the learner’s own standard of knowledge. This depends upon the student’s critical thinking skills, and what conceptual constructions he/she will accept as good enough to pass the tests of being logical, reasonable, and useful in making predictions (von Glasersfeld, 1988). It is helpful and appropriate here to use as a model for concept acquisition the theoretical framework for conceptual change which has been researched and proposed by Posner, Strike, Hewson and Gertzog (1982) and by Strike and Posner (1985), which draws heavily on contemporary Philosophy of Science. The theory views learning in general as a process of conceptual change, which is in keeping with a constructivist view. New phenomena which fit the learner’s existing framework are assimilated, and those which require reorganizing or replacement of the existing framework are accommodated. All new information has to be interpreted in terms of the learner’s existing conceptual framework. In order to change that framework to accommodate new or different concepts, the authors theorize that four conditions seem necessary:

1. There must be dissatisfaction with existing conceptions. That is, anomalies must be sufficiently bothersome to justify the effort of change.
2. The new conception must be understandable, initially perhaps through metaphor or analogy.
3. The new conception must appear likely to solve the unsolved problems.
4. The new conception should show promise of directing the further construction of new knowledge.

The influence of Kuhn’s (1962) theory of a prevailing paradigm and scientific revolution is evident here, as well as Lakatos’s (1970) theory of research programs, including the protected “hard core” concepts, and the protective “outer belt” concepts.

The conceptual ecology of the learner is crucial in determining when the above conditions will be met (as mentioned under the sixth source of misconceptions, and is often problematic with science students. Osborne and Wittrock (1985) and von Glasersfeld (1988) pointed out that knowledge restructuring takes place when the newly structured knowledge passes tests to the satisfaction of the learner. That is, for misconception-free concept acquisition to occur, the learner must recognize and refuse to accept anomalies. S/he has to have rigorous standards of
and sources of misconceptions and lack of understanding. Yager cited to understanding of scientific concepts. Faced with literature, there is a basic dilemma here in that acquiring these attitudes is an important part of the learning process, and yet learning Science does depend on having these attitudes. This having been recognized, it is natural and inevitable that the learning process will include a continuous series of acquiring and changing perceptions.

SIGNIFICANCE OF THE PROBLEM

Many misconceptions in science, particularly in Chemistry, are being propagated unknowingly by teachers and textbooks, in a variety of ways. Typical classroom practices make the acquisition of misconceptions very nearly inevitable. Language is undoubtedly the single most problematic source of misunderstanding, with teachers using terminology with one meaning in mind, and students interpreting the same terms very differently (Jacobs, 1989; Veiga, Pereira and Maskill, 1989). In addition, students frequently do not have the prior knowledge needed to acquire new scientific concepts (Champagne, Klopfer, and Anderson, 1980; Linn, 1980; Sternberg, 1981; Clement, 1983; Glaser, 1984; Green, McCloskey, and Caramazza, 1985). Teachers commonly feel great time pressure, and often cover large quantities of material quickly, overlooking the limits of the processing capacity of students (Johnstone, 1984; Sweller, 1988). The mismatch of student cognitive level with the cognitive demands of the material to be learned is a common occurrence in science subject areas involving any extensive abstract reasoning (Linn, 1982; Lawson, 1985). This practice invites errors in reasoning which in turn lead to the development of misconceptions (Case, 1975; Chiapetta and McBride, 1980; Andersson, 1986).

All these factors contribute to a learning situation in which it is likely that students will often be unable to process information sufficiently to construct an acceptable level of understanding of scientific concepts. Faced with the enormity of the task, students often resort to guessing, memorizing, and low standards of logic (Johnstone, 1984). Furthermore, studies cited in the literature search indicate that when teachers are asked to predict their students' understanding and test performance, the resulting predictions are very inaccurate (Ivowi, 1986; Yager and Penick, 1987). This indicates that the teachers are not fully aware of students' misconceptions and lack of understanding prior to testing them. Articulation of the nature and sources of students' misconceptions in Chemistry should assist teachers in designing instruction and assessment to better match the students' learning process, and prevent many currently widespread misconceptions.
LIMITATIONS OF THE STUDY

This study is an ambitious effort to give a synopsis of the cognitive difficulties which students of Chemistry typically encounter. While this should be beneficial to teachers of Chemistry, the scope of the study inevitably involves certain shortcomings.

The greatest limitation of the study is that the interview methodology employed to explore student concept formation is not designed specifically to detect all the error sources described. This applies in particular to Overtaxing the Short Term Memory of the Student, and Mismatching the Cognitive Demands of the Material to the Level of Cognitive Development of the Student. While the short term memory demands of the learning task are measureable, this measurement is quite complex and would not be feasible in this study. The study could, however, be strengthened by testing the subjects for level of cognitive development. It was not done in this study because of severe time pressure. The time requirement on the subjects was quite considerable, and it would have been virtually impossible to schedule an additional testing session.

Another weakness of the study is the assumption that the subjects' learning was not affected by other non-cognitive factors, such as time pressure, anxiety, social pressure, personal motivation, etc. This, of course, could not be the case. Insofar as could be known under the circumstances of the study, students with very unusual situations were avoided as subjects.

The study is also limited by the fact that the sample was composed of volunteer subjects. This was a necessary condition due to the demands which the study placed on the subjects.

OVERVIEW OF THE STUDY

The primary goal of this study is to articulate and substantiate a classification system of instructionally related misconceptions characteristic of concept acquisition in introductory Chemistry, using a unit on chemical equilibrium as the study vehicle. The methodology had three distinct phases: First, the synthesis of historical research leading to the formulation of the classification; second, the gathering of qualitative data through an interview process; third, the quantitative analysis of that data. The fundamental nature of the study is qualitative, and as such is not designed to prove hypotheses. However, the study was designed to examine the validity of the following hypotheses:

\[ H_1: \text{ Misconceptions in the learning of Chemistry can be categorized as particular types.} \]

\[ H_2: \text{ Misconceptions in the learning of Chemistry result primarily from the instructional process.} \]

\[ H_3: \text{ Misconceptions in the learning of Chemistry do not arise from} \]
experiential misconceptions, i.e., longstanding deeply
imbedded misconceptions resulting from real life experiences.

H4: The instrument of measurement, the probing interview, has a
positive effect on the development of understanding in
Chemistry.

The above hypotheses were investigated both quantitatively and qualitatively as described below.

METHODOLOGY

This study observed a group of high school Chemistry students as they covered a unit on chemical equilibrium, using weekly interviews to assess the development of their understanding of equilibrium concepts. The unit included general principles of reaction equilibrium, acid/base equilibrium, and solubility equilibrium. The chemical equilibrium unit was taught by a teacher who had ten years of classroom experience. This teacher also has Master's degrees in both Chemistry and Education, and is highly involved in professional development.

The interview protocols used were generated based on classroom observations carried out during the course of instruction on the unit, and on analysis of the textbook materials. Four or five tape recorded interviews were conducted with each subject over the duration of the unit. The tape recorded interviews were transcribed verbatim, and student errors were coded based on the proposed classification scheme developed based on the researched literature, and refined in two previous pilot studies (See Appendix). The coding was first done independently both by an experienced Chemistry teacher and by the investigator, and then re-done together to reach consensus on discrepancies. The data generated by the coding process was analyzed both quantitatively and qualitatively.

Sample

A volunteer sample of 12 students was drawn from a population of 48 college preparatory students in three sections of the same course in college preparatory Chemistry at an independent boarding school in northeastern Massachusetts. This group consisted of 6 males and 6 females, 6 sophomores and 6 juniors. Their PSAT scores ranged from 68 to 114, with a mean of 98, and prior achievement in the Chemistry course, as measured by their first semester grade (reported as a percentage), ranged from 72 to 91, with a mean of 81. The subjects in this sample were matched as closely as possible with another student in the population by gender, PSAT, and prior achievement in Chemistry for the purpose of control and comparison of post unit achievement. The results of ANOVAs performed on the characteristics of the students in the two groups showed no significant differences between the groups at the start of the study.
Development of the Codes

The coding schema consists of seven principal categories of sources of error, each of which is designated by a multiple of ten. An eighth category was used as a utility category, to record other phenomena in the interview process. Category 80 is not an error category. The categories were as follows:

10 Language related misunderstanding and misinterpretation.
20 Deficient prior knowledge on the part of the learner.
30 Overtaxing the short term memory of the learner.
40 Mismatch of cognitive demands of the subject matter with the level of cognitive development of the learner.
50 Choice of mental strategy inappropriate to the subject matter on the part of the learner.
60 Low standard of epistemology on the part of the learner.
70 Experiential misconceptions.
80 Other coded phenomena.

Each of categories 10 through 70 contains from one to nine subcategories, which are specific types of errors within the category, which were drawn from the literature as well as the investigator’s career classroom experience. A detailed description of the final coding scheme is found in the appendix.

The categories were developed based on the literature search, and did not change over the pilot studies or the principal study. However, the number of subcategories was expanded in the course of the coding process, both in the pilot studies and in the principal study. The expansion was necessitated by the identification of errors which seemed to belong to one of the main categories, but a specific description of what was occurring did not fit the existing subcategories. When this occurred in the principal study, both raters worked together to describe the error, and the description was then added as a subcategory under the appropriate category.

Content validation of the categories was initially established informally by frequent discussion with three experienced Chemistry teachers as the codes were developed and refined. Final content validation was established by submitting the classification system to two other experienced teachers of Chemistry for comment and critique after the codes were finalized at the end of the study. Both experts judged the classification to be a comprehensive summary of sources of error in learning Chemistry.

The Study

The course being pursued by the subjects was a typical college preparatory Chemistry course, using the text Chemistry: A Modern Course, by Smoot, Smith and Price, 1990, a Merrill
Publishing textbook which is widely used throughout the United States. During the first semester, the students studied chemical properties, formulas, moles, equations, and stoichiometric relationships involved in chemical reactions, as well as gas laws and kinetic molecular theory. During the second semester, the students studied atomic structure, periodic properties of the elements, chemical bonding, and the properties of solutions. Rates and mechanisms of reactions were being studied during the first two weeks of the investigator’s observations. This topic led directly into the unit on chemical equilibrium, covered in chapters 23-25 of the text, which was the topic of investigation in this study.

Procedure

Class Observations

One class (of the three equivalent sections) was observed by the investigator for each lesson in the unit. The observation schedule was set up to cycle through the three classes consecutively, repeating the cycle throughout the observation period. During the class observations, the investigator sat at a lab table, directly adjacent to the lecture area of the room. From this vantage point, all of the students faces, as well as a full view of the teacher, was possible. The proceedings of each class were tape recorded, as the investigator took notes. The investigator’s notes included the content being covered, questions asked by the teacher, students’ responses, questions asked by the students, and comments on the proceedings.

As soon as possible following the class, the investigator listened to the tape of the class proceedings, filled in the notes on the class, and checked off possible sources of error to look for in the interviews. This follow-up listening always took place within the same day as the original observation.

Text Analysis

The primary purpose of the text analysis was to determine the content covered in the homework assignments, in order to incorporate that content into the interview protocols. The secondary purpose was to anticipate error sources which might be caused by the text. The investigator took notes on the text, and also used an observation checksheet to record instances of possibly misleading presentation of material. The information obtained made possible a more probing set of interview protocols.

The Interview Protocols

Five interview protocols were written by the investigator, based on the data collected as described above. The protocols include the primary questions asked of each subject, as well as the most frequently asked probe questions. The probe questions are follow-through questions asked by the investigator when a student gave an unclear or incomplete answer to the primary question, or if a student gave no answer to the primary question after 5 seconds. Any laboratory
materials required for a particular interview are listed at the head of that protocol. The topics of the five protocols are as follows:

1. Prior background on the concept of equilibrium and vocabulary used in the unit.
2. Qualitative aspects of equilibrium: Reaction reversibility and LeChatelier’s Principle.
3. The nature of the Equilibrium Law, the meaning of the equilibrium constant, \( K_E \), and the concepts of \( K_A \) and \( K_W \).
4. Strong vs. weak acids (and bases), and the use of \( K_A \) and \( K_W \).
5. Solubility equilibria.

The complete texts of the five protocols, including probe questions, are available from the author.

**The Interview Procedure**

The interviews took place in a quiet study room located adjacent to the Chemistry classroom/lab. The investigator posed each question of the protocol, in order. If the subject’s response did not fully reveal his/her understanding of the question, further probe questions were used. If a subject did not respond immediately, a silence of five seconds was allowed before further comment or probe question by the interviewer. If it became apparent that the subject could not answer the question, even with probe questions, the investigator summarized the answer in order to be able to go on with the interview. This was used as a “Stop Mechanism,” and was coded as that.

**Transcription**

The interviews were transcribed verbatim by the investigator, using a transcription tape recorder. A single spaced format was used, allowing a wide right hand margin, to accommodate the codes. Eighty percent of the transcription was completed in the summer immediately following the study. The remaining twenty percent was transcribed in the following six months, during the school year.

**The Coding Procedure**

The coding of the transcribed interviews was done by the investigator and an experienced Chemistry teacher. The first phase of training for the teacher-coder involved her reading the overview of the study and the literature review, upon which the categorization of error sources was based. Discussion then took place with the investigator to clarify any questions regarding the meaning of the codes. A good part of this discussion involved example errors, and how they would fit into the categorization schema.

The second phase of training involved the coder and the investigator together, coding two interviews from a previous pilot study. The coding involved reading through the transcript
to identify errors. When an error was identified, it was underlined in the transcript text. A decision was then made on the principal category to which the error belonged. Finally, a code was assigned for the error description within the chosen category which best fit the identified error. This served to further clarify how the codes were to be applied.

Finally, the coder and the investigator coded a sample interview separately, using the method described above, and then re-coded it together to establish consensus. This process was repeated for a second interview. Although there was agreement of only approximately 80% at this point, consensus sessions resolved all differences of opinion between the coder and the investigator. Therefore, the coding of the interviews from the study was started.

The interviews were coded as complete sets (all Interview 1, all Interview 2, etc.) separately by the coder and by the investigator, who each generated a set of codes for each interview directly on the transcript sheets. These were then recorded in a summary sheet. The coder and investigator then went through each set of codes together, noting differences. They then went through each entire interview together, affirming agreements, and discussing differences, until consensus was reached on all codes. Where changes were made in the original choice of code by both the coder and the investigator, the change was recorded on the original transcript sheet in green, providing a permanent record of those changes. In several cases, new subcategories were described jointly by the coder and the investigator, resulting in the final version of the codes reported in the appendix.

DATA ANALYSIS

First of all, the data was collapsed from the separate subcategories into the seven major categories for purposes of analysis. The rationale was that the subcategories were used to identify what seemed to be going on with the subject, and to guide and justify the selection of principal category. The frequency of coding instances for the seven principal categories is found in Table 2.
Finally, a further collapse of the data was carried out, with the following rationale. The frequency distributions of the error categories, found in Table 2, suggested that the occurrence of Category 7 errors, Experiential Misconceptions, was too small to be significant. Therefore, this category was eliminated as a contributor to errors in concept formation in Chemistry. Categories 30, 40, 50 and 60 were collapsed into one single category, *All Other Cognitive Error Sources*, because they each involve a cognitive process which, while recognized in the literature, were not specifically detectable by the methodology used in this study. While each of these four factors is important in the learning process, and while an experienced teacher can have a sense that a particular category is causing difficulty, the coder can only infer that any of these categories is operative in a given case. The result of this final collapse of the data is found in Table 3.
This collapsing of the data gives a realistic representation of the sources of error in concept formation in Chemistry:

1. Errors involving Language.
2. Errors involving Prior Knowledge
3. All other cognitive errors: Errors involving Short Term Memory, Cognitive Development, Mental Strategies, and Standards of Epistemology.

**Analysis of Reliability of Coding of Errors**

A detailed analysis was carried out to evaluate the reliability of coding both between interviews for each coder, and between coders. This analysis is too lengthy to include in this paper, but is available from the author. Agreement between coders was approximately 80% prior to consensus sessions, so it is clear that consensus sessions are needed to establish a final distribution of errors for a set of interviews. However, this was both expected by the investigator and predicted by the expert validators, due to the size and complexity of the classification system.

**The Set of Error Codes**

**Frequency of Occurrence of Error Sources**

The lack of *Experiential Misconceptions* as a source of error is notable. Although this source of error is reported in the literature as a serious cause of difficulty in concept formation, it was a basic premise of this study that experiential misconceptions are not important as a source of error formation in Chemistry. The frequency data lends support to this premise.

The total error distribution was found to be roughly equally divided among the three major categories, with 37% of detected error attributable to *Language* related difficulties, 32% of detected error attributable to the *Prior Knowledge* of the student, 30% of detected error

---

**Table 3**

Frequency of occurrence of errors by collapsed categories

<table>
<thead>
<tr>
<th>Interview #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Error Source</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>67</td>
<td>69</td>
<td>74</td>
<td>116</td>
<td>34</td>
<td>360</td>
</tr>
<tr>
<td>Prior Knowledge</td>
<td>33</td>
<td>30</td>
<td>102</td>
<td>112</td>
<td>47</td>
<td>324</td>
</tr>
<tr>
<td>Other Cognitive Factors</td>
<td>13</td>
<td>45</td>
<td>78</td>
<td>109</td>
<td>55</td>
<td>300</td>
</tr>
<tr>
<td>Experiential</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>15</td>
</tr>
</tbody>
</table>

This collapsing of the data gives a realistic representation of the sources of error in concept formation in Chemistry:
attributable to Other Cognitive Factors, and only 1% of detected errors being attributable to the Experiential Misconceptions of the learner.

The distribution of error sources varies across individual interviews, however. This is an expected result, since the nature of the content of the interview protocols varies. (Interviews 1 and 2 were primarily qualitative, involving verbal explanations. Interviews 3 and 4 involved some verbal interpretations of problem situations, but involved considerably more abstract thinking, and placed a greater emphasis on mathematical applications. Interview 5 involved a fairly equal mixture of quantitative and qualitative applications.)

**Analysis of Errors by Interview**

Table 4 presents the results of a one way ANOVA between the five interviews, comparing sources of error detected. The analysis indicates that there are significant differences in the types of errors detected over the five interviews for Language, Prior Knowledge, and Other Cognitive Factors. No significant differences among the interviews was found for Experiential errors.

**Table 4**
F Ratios for one way ANOVA on interview number and category of error.

<table>
<thead>
<tr>
<th>Error Source</th>
<th>df₁</th>
<th>df₂</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>4</td>
<td>49</td>
<td>3.37</td>
<td>&lt;.02</td>
</tr>
<tr>
<td>Prior Knowledge</td>
<td>4</td>
<td>49</td>
<td>7.07</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Other Cognitive Factors</td>
<td>4</td>
<td>49</td>
<td>5.46</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Experiential</td>
<td>4</td>
<td>49</td>
<td>.85</td>
<td>&gt;.05</td>
</tr>
</tbody>
</table>

The results of a post hoc comparison using the Fisher protected t-test are presented in Table 5. The results of this analysis indicate that there are differences in types of error sources when the content of the material changes in nature, as expected from the above discussion of content differences from interview to interview.

In the category of Language, there were fewer differences between interviews than expected. Differences between Interview 2 and Interview 4, as well as between Interview 4 and Interview 5 are probably due to the higher verbal demands of Interview 2 coupled with the
higher mathematical demands of Interview 4. Differences between Interviews 3 and 4 were unexpected. However, the material in Interview 4 was the most challenging of all the five sessions, and errors of all categories peaked in this session. The very high cognitive load posed by Interview 4
Table 5
Post hoc analysis of interview number versus error source, using the Fisher Protected t-Test

<table>
<thead>
<tr>
<th>Language</th>
<th>Interview</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td>&lt;.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td>&lt;.05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prior Knowledge</th>
<th>Interview</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td></td>
<td>&lt;.05</td>
<td>&lt;.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td>&lt;.05</td>
<td>&lt;.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td>&lt;.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td>&lt;.05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other Cognitive Factors</th>
<th>Interview</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td></td>
<td>&lt;.05</td>
<td>&lt;.05</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td>&lt;.05</td>
<td>&lt;.05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

did appear to trigger errors which subjects would not ordinarily commit. The differences between Interview 4 and Interview 5 are probably due to the content differences of the protocols, with Interview 5 having a higher verbal component, and considerably less mathematics.

In the category of Prior Knowledge, all differences across interviews are expected, due to differences in protocol content. Significant differences are found between Interview 1, which has a high verbal load, and Interviews 3 and 4, which have a high mathematical load. The same differences are found between Interview 2, which is high verbal, with Interviews 3 and 4. Significant differences are also found between both Interviews 3 and 4 with Interview 5, presumably due to the lesser mathematical demands of Interview 5 as compared with Interviews 3 and 4. Of the subcategories of Prior Knowledge which involve the student's mathematical background, (See #24 through #27 in appendix) 75% of the coded errors were in Interviews 3 and 4, 20% in Interview 5, and only 5% in Interviews 1 and 2.
In the category of Other Cognitive Factors, all detected differences across interviews are expected, again due to the different nature of the content of the protocols. The fact that there is a significant difference between interview 1 and Interview 5, but not between Interview 2 and Interview 5 is most likely due to the fact that Interview 2 applied LeChatelier’s Principle to general equilibrium systems, whereas Interview 5 required the application of LeChatelier’s Principle to solubility systems.

In summary, the experimental data shows the system of categorization to be a comprehensive summary of the errors which occur in student concept acquisition in Chemical Equilibrium. While the errors do represent those which are problematic for students, the distribution of errors over the categories appears to be a function of the demands of the particular topic being studied.

Table 6
Descriptive statistics for PSAT scores, semester 1, and semester 2 Chemistry grades for experimental, matched control, and unmatched control groups.

<table>
<thead>
<tr>
<th>Group:</th>
<th>Experimental</th>
<th>Matched Control</th>
<th>Unmatched Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Semester 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>80.9</td>
<td>80.3</td>
<td>80.3</td>
</tr>
<tr>
<td>S. D.</td>
<td>6.83</td>
<td>6.88</td>
<td>5.66</td>
</tr>
<tr>
<td>Max.</td>
<td>91</td>
<td>93</td>
<td>87</td>
</tr>
<tr>
<td>Min.</td>
<td>72</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Range</td>
<td>19</td>
<td>23</td>
<td>17</td>
</tr>
<tr>
<td>Semester 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>80.58</td>
<td>80.75</td>
<td>78.92</td>
</tr>
<tr>
<td>S. D.</td>
<td>7.88</td>
<td>6.74</td>
<td>6.96</td>
</tr>
<tr>
<td>Max.</td>
<td>94</td>
<td>90</td>
<td>89</td>
</tr>
<tr>
<td>Min.</td>
<td>71</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Range</td>
<td>23</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Equilibrium Unit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>81.17</td>
<td>75.17</td>
<td>75.42</td>
</tr>
<tr>
<td>S. D.</td>
<td>8.85</td>
<td>9.49</td>
<td>10.65</td>
</tr>
<tr>
<td>Max.</td>
<td>95</td>
<td>85</td>
<td>91</td>
</tr>
<tr>
<td>Min.</td>
<td>68</td>
<td>56</td>
<td>51</td>
</tr>
<tr>
<td>Range</td>
<td>27</td>
<td>29</td>
<td>40</td>
</tr>
</tbody>
</table>

The subjects’ actual test scores, as a group, show a six point advantage in favor of the interviewed group. It is not possible to determine, definitively, the cause of the
difference in grades. The most obvious possible cause is the additional six hours of time spent by each subject in processing the material to be learned, during the course of his or her interviews. The interview process was also necessarily designed to give students feedback on incorrect concept formation as the study proceeded. Not least of all, each subject, knowing that s/he would be interviewed individually on material being covered in the course, should certainly have felt motivated to keep up with the work at hand. In summary, there are other variables affecting the students' performance on the equilibrium unit which are not controlled for in this study. Nor was it the purpose of this study to determine the effect of the interview process. However, it is an interesting outcome to examine. There appears to be a small positive effect associated with the interview process.
References


Appendix: Coding Scheme for Interviews with Examples

It is the purpose of this appendix to clarify the meanings of the subcategories to a potential coder. A potential coder needs first to read the literature research section, in order to become familiar with the research background suggesting the coding scheme. Examples in the section below are drawn from the literature references, the pilot studies, and the investigator’s classroom experience.

It is a basic assumption of this research that the coders must be experienced Chemistry teachers as a basic prerequisite to recognizing errors encountered in the interviews. A major difficulty for the coders in this study proved to be confusion between the subcategories of the categories of Language and of Prior Knowledge. It was helpful and effective to determine the principle category first, and then to select the subcategory which best fit the error.

The following outline lists typical examples of the error sources listed as a subcategories of each error type.

10 Language related misunderstandings

11. Error related to using the everyday meaning for words which have a context-specific meaning in science.
   The use of the word “significant” figure to describe a measured digit in an experiment or the use of the word “spontaneous” to describe a reaction which is thermodynamically possible both can cause confusion and misunderstanding. Students tend to attach the meanings that they already know to these words, rather than the newly learned scientific definition.

12. Technical words being defined technically, and used without understanding.
   The use of any technical word by a student, but particularly those terms which have been recently introduced in the course, warrants verification of meaning in this study. The use of such terms in appropriate contexts frequently masks a lack of understanding. Examples typical of this unit of study include “strong”, “weak”, and “buffer” as applied to acid/base solutions.

13. Using a word with a technical meaning inappropriately when applying to specific situation or phenomenon.
   This subcategory refers to a student’s use of a technical term in an inappropriate context. For example, in one pilot interview, a subject used the “law of conservation of matter” in explaining aspects of the “law of mass action.” Further questioning revealed that his conceptualization of the situation was completely irrelevant.
14. Misunderstanding or misuse of a symbol.
An error of this type includes two situations:
  a. The student does not know the verbal or conceptual entity which a symbol represents.
     A typical example in this unit was student failure to recognize or use correctly the square bracket symbol, [   ], to represent concentration in molarity.
  b. A student confuses one symbol with another.
     A typical example of this is the letter “m.” In small case, m, it can be used for the prefix milli, the unit meter, the measurement mass, and the concentration unit of molality. In large case, M, it is used for the concentration unit of molarity, but also for molecular weight in some Ideal Gas Law calculations. As a result, students sometimes confuse the different possible meanings.

15. Overextension of an analogy.
This can occur when a student extends the use of an analogy to a situation in which it does not apply.
For example, there is a demonstration used to exemplify equilibrium in which a hole is punched in a one liter plastic soda bottle. Water is allowed to flow into the top of the bottle. After a minute or so, the water flow is adjusted so that the level of water in the bottle remains the same, i.e., the rate of flow of water into the bottle equals the rate of flow out of the bottle. The problem with this analogy for equilibrium is that the same water which is entering the bottle is not the same water which is leaving the bottle. This could not be the case in chemical equilibrium, which requires a closed system. But the analogy is useful in conveying the idea that the same amount of material is in a particular state or condition in an equilibrium system, and also that that quantity of material is not the same actual atoms or molecules, but only the same overall amount.

16. General misinterpretation or misunderstanding of the overall statement or question.
In some cases, it is clear that the subject does not comprehend the overall meaning of a question or statement, perhaps because of the wording. The student may answer an unintended question, state that s/he does not understand the question, or may ask for clarification of a term or phrase in the statement. If rewording of the question or statement (by the interviewer)
provides clarification, a #16 should be coded.

17. Inability to relate language to its graphical representation.
If a student does not understand a graph or a diagram, the interviewer should ask him/her to read and explain the labels of the axes, and the accompanying legend. This type of probe question will clarify if the source of misunderstanding is language related. If the problem lies with understanding what it is that has been plotted, #17 is the appropriate code.

18. Incomplete definition, not clearly specified.
This can generally be described as imprecise use of language. This code is used when the subject gives an answer which has incomplete detail. For example, in describing the dissolving process, the subject may refer to the role of the solute only, without referring to the solvent, or describe molarity as simply “moles over liters”, without any specification of units. Or a subject may describe a base as “the opposite of an acid.” The interviewer should ask for clarification. Code #18 should be used when the subject is unable to fill in the details.

19. Inability of a student to articulate his/her thoughts.
This code is used when a subject makes such comments as, “I understand, but I just can’t explain it,” or “You know what I mean,” without being able to provide specific follow-up. This subcategory frequently is evident when a subject has just completed a problem more or less successfully, but has difficulty putting the steps or reasoning followed into words.

20. Misunderstandings related to prior knowledge of learner

Note: Potential coders need to read through the content outlines of the chapters which have been covered in the text, in order to know what background information students should have.

21. Chemistry fact or vocabulary word not known to the subject.
This category is self explanatory, except to specify that it refers to individual facts or words, such as molarity, and not entire concepts.

22. Chemistry fact or vocabulary used incorrectly or incompletely by the subject.
A common example encountered in this unit was the misuse of the term molarity of solutions. Students were observed to use moles alone, without dividing by volume, as a molarity, as well as to use molarity directly as moles, without using the volume to convert to moles from the concentration unit.

23. Chemical procedure misused by student.
A procedure is defined here as a series of steps used to carry out a typical chemical problem. Examples which were encountered include:

a. Finding molarity by dividing grams of solute by volume of solution.

b. Given the quantity of one substance in a balanced equation, students committed various errors in finding the required quantity of another reactant or product in the equation: Failing to convert grams to moles, incorrect use of the equation ratio, finding the wrong substance, etc.

24. **Algebraic error: Mechanics.**
   Incorrect solution of a correctly set-up algebraic equation. Examples include multiplying instead of dividing, finding the reciprocal of the answer, etc.

25. **Algebraic error: Reasoning.**
   Failure to set up an algebraic expression correctly. For example, students were observed to set up incorrect quantitative ratio between two different species in a balanced equation.

26. **Algebra fact or definition not known or misused.**
   This is coded when the algebraic content required for a situation is not known or not recognized by a student. Examples would include not knowing the equation for a straight line relationship, not knowing the meaning of a logarithm, or not knowing how to use a logarithm.

27. **Improper use of calculator by student.**
   Insofar as possible, the interview transcripts include all student comments which indicated a problem with producing correct answers from the calculator. Therefore the coding of these errors is reasonably straightforward. Working back from students’ incorrect answers is another method of determining if this code should be used.

28. **Confusing a newly introduced concept with a previously learned concept.**
   An example of this was found in discussing chemical equilibrium. Several students confused the visible changes in systems undergoing a LeChatelier shift with “physical changes.”

29. **Previously covered material has been forgotten or confused.**
   This would include a subject’s lack of sufficient familiarity with any fact, concept, or procedure previously covered in the course.
   Notable examples in the unit of study investigated were the use of electron dot formulas to represent molecules, stoichiometric relationships, formula writing, and the concept of molarity.
30. Error due to overtaxing the short term memory of the learner

The documentation of errors due to overtaxing the short term memory of the subject is problematic in this research, because the procedures employed are not specifically designed to detect this source of error. However, when analyzing an interview, a knowledgeable Chemistry teacher is likely to recognize when a student has not sufficiently processed a chemical procedure so that the procedure functions as a single unit in short term memory. Thus mental overload results. In this study, no further breakdown of this category was practical. It was in evidence in a number of the interviews documented in the “Discussion” section of Chapter 5.

31. Misunderstanding caused by too rapid coverage.

For example, in setting up a chart of equilibrium concentrations to map the progress of a reaction as equilibrium is established, the student will need to consider the following aspects of the system: the balanced equation, the stoichiometric relationship between the separate compounds in the equation, and the concept of molarity. If the student has not sufficiently automated each of these procedures previously, the short term memory load as s/he proceeds through the problem solution will be too great.

40. Error due to mismatch of cognitive demands of the subject matter with cognitive developmental level of the learner

Although research has established this to be an important source of conceptual errors in learning Chemistry, it was not expected to be evident in the interview format used. However, it does appear to be operative in a number of the interviews. For example, subjects sometimes gave very literal or concrete answers to abstract questions, and #41 was coded in these cases.

41. Subject matter is too abstract or formal for the developmental level of the student.

A good example is the common student misunderstanding of the role of spectator ions. Students frequently ask why spectator ions should not be in an equation if it is necessary that they are present in a reaction container.

50. Error due use of inappropriate mental strategies.

51. Use of algorithms without understanding.

For example, Students quickly pick up the algorithm that $\text{pH} + \text{pOH} = 14$. But probing will sometimes (or often) reveal that they do not connect the numbers with the water equilibrium or with actual concentrations of hydroxide and hydronium ions.

The teaching of formula writing by having students “criss-cross” the charges
of the ions involved can produce correct formulas quickly, but without any understanding on the part of students as to why the compound actually has that particular ratio of ions.

52. **Rote learning of material not yet understood.**

In studying acid/base Chemistry, students memorize the water equilibrium expression, $K_w$, and its value, and can often solve “worksheet” problems quickly and easily. But they frequently do not recognize the same concept in other contexts, because they have basically just memorized the information. In another example, when a subject was probed as to why he left water out of a particular equilibrium expression, he said, “Because it doesn’t matter.” He actually had no understanding of the reason.

53. **Error in logic (including getting right answer for wrong reason.)**

A logical error observed repeatedly involved the failure to take into account that when either $[\text{H}_3\text{O}^+]$ or $[\text{OH}^-]$ changes, the other must also change in order to obey the $K_w$ constant value. When faced with the contradiction, students were generally more willing to allow the value of the constant to change, rather than either of the concentrations in question.

60. **Errors due to insufficiently rigorous standard of knowledge on the part of the learner.**

Although each of the subcategories below is distinguishable from the subcategories under “Mental Strategies,” a logical argument could be made for collapsing the “50’s” and “60’s” into one category. They were treated under this separate major category in this study.

61. **Guessing.**

This subcategory is intended to be used to identify straightforward and intentional guessing only. If a subject has been following a problem or discussion closely, and attempting to think it through logically, #61 should not be coded. Guessing should not be confused with a subject’s uncertainty or tentativeness, usually marked by a raising of the voice at the end of an answer or statement.

62. **Tolerance of illogical statements or conclusions (Learner is generally aware of the error.)**

In discussing a problem involving a buffered solution containing 0.15 M HCN and 0.10 M NaCN, a subject wrote out the equation, explained what was in solution, and when substituting into the equilibrium expression for the equation:
HCN + H₂O ⇄ H₃O⁺ + CN⁻

just quickly said that [H₃O⁺] and [CN⁻] would have to be the same. It was obvious from the context that the subject knew that was not true in this case, but was willing to drop it at this point, rather than think through the implications.

63. **Insufficient scrutiny of answer or conclusion** (Learner is generally not aware of error, but is expected to be.)

When asked the number of chemical bonds typically formed by the families on the periodic table, one subject started from the right, with 0,1,2,3, (starting with the Noble Gases) and then switched to start from left. She then said, “So it must be 0,1,2,3, for the other side also,” but this time starting from the Alkali Metals. It was something she should have known so automatically that she did not bother to think if her answer made any sense.

70. **Experiential misconceptions.**

This section involves a very different type of error, not expected to be commonly found in Chemistry, since chemical phenomena not are occurring on a macroscopic level. This category includes any error or misconception formed by misinterpretation of everyday phenomena which can be detected by the senses. It was not intended to make any further subdivision of this category.

71. **Experiential misconception detected.**

72. **Error made due to existing experiential misconception.**

73. **Direct correction of experiential misconception.**

80. **Detection of facilitating of error correction.**

This category is not an error category, but rather a record keeping category, which was used to record occurrences of investigator actions and questions.

81. **Probe question used to cue or encourage an answer from the subject.**

82. **Interviewer summarizing, giving correct answer/ interpretation.**

This procedure was necessary in order to facilitate getting past a concept which the subject clearly did not know. It is essentially a **STOP mechanism.**

83. **Student self correction.** (Not including simple mis-speaks.)

84. **Student makes same error after previous correction by interviewer**