One of the four algebra word problem structures found in K-12 textbooks is the propositional relation structure (Maye, 1982). This type of problem asks students to establish equivalences between the variables or noun referents in the problem. The literature available indicates that students have inordinate difficulties when trying to solve a propositional relation type of problem. The literature says little about how key context features are assigned to the text of the algebra word problem and how these features affect student performance or preference of algebra problems in general. Three key context features of algebra word problems were identified: familiarity, imageability, and variable type (discrete or continuous). These three key contextual features have been shown to have performance effects on arithmetic and algebra problems. The analysis in this paper examines these key context features of algebra problems in relationship to pictorial, symbolic, and verbal representational formats. In addition, an information processing model of the algebra word problem solving process is presented. The features in which key contextual features may influence problem processing and the construction of solutions is discussed. This paper supplies a rationale, definitions, and a procedure for assigning key context features to an algebra word problem of the propositional relation type. These procedures were used to construct a set of 16 algebra word problems that systematically varied the key contextual features identified. These problems were proposed to be used to study several variables that may be important to understanding the nature of difficulties students face when solving a problem with a propositional relation structure. The model developed and presented could be generalized to other algebra problems of different structures or other domains. (Contains 45 references.) (MDH)
"KEY CONTEXTUAL FEATURES OF ALGEBRA WORD PROBLEMS:
A THEORETICAL MODEL AND REVIEW OF THE LITERATURE"

Ramzi Nasser, UMASS Lowell
James Carifio, UMASS Lowell

Paper presented at the Annual Conference of the Eastern Educational Research
Association, Clearwater, Florida, February 17-22, 1993

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

James Carifio

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE
"Key Contextual Features of Algebra Word Problems: A Theoretical Model and Review of the Literature"

Ramzi Nasser, UMASS Lowell
James Carifio, UMASS Lowell

ABSTRACT

One of the four algebra word problem structures found in K-12 textbooks is the propositional relation structure (Mayer, 1982). This type of problem asks students to establish equivalences between the variables or noun referents in the problem.

The literature available indicates that students have inordinate difficulties, when trying to solve a propositional relation type of problem (Sims-Knight & Kaput, 1983a). This particular literature says little about how key context features are assigned to the text of the algebra word problem and how these features affect student performance for preference of algebra problems in general.

We have identified several key context features of algebra word problems based on a review of the literature in this area. These features are familiarity, imageability and variable type (discrete or continuous). These three key contextual features have been shown to have performance effects on arithmetic and algebra problems (e.g., Sims-Knight & Kaput, 1983a & 1983b; Lyda & Franzen, 1945; Sutherland, 1942; Brownell & Stretch, 1931; Washburne & Osborne, 1926 and Horwitz, 1980). The analysis in this paper examines these key context features of algebra problems in relationship to different presentations and responding formats; namely, pictorial, symbolic and verbal format. An information processing model of the algebra word problem solving process is presented which is based in part on Kintich and Greeno (1985) template and story-line framework. The ways in which key contextual features may influence problem processing and the construction of solutions is also discussed.

This paper supplies a rationale, definitions and a procedure for assigning key context features (i.e., clothing of the problem structure) to an algebra word problem of the propositional relation type. These procedures were used to construct a set of 16 algebra word problems which systematically varied the key contextual features identified. These problems will be used to study several variables that may be important to understanding the nature of difficulties students face when solving a problem with a propositional relation structure. The model developed and presented could be generalized to other algebra problems of different structures or other domains.

The major outcomes of this study is a method to assign key context features to a set of domain referenced algebra problems, which could be generalizable to other algebra problems of different structures.

Overview

This paper provides a rational, definitions and a procedure for assigning key contextual features to the generic structure of algebra word problems. The key context features of algebra word problems are the "clothing" of the generic structure of a given type of algebra word problem. Given these points we surveyed the literature on the types of algebra word problems generic structures that solvers could encounter. Among the few studies we found (e.g., Paige & Simon, 1966 and Riley, Greeno, & Heller, 1983), Mayer's (1982) study was most helpful.

In his review of the K-12 textbooks, Mayer developed a general model and set of descriptions of types of algebra word problems. In terms of general models of algebra word problems, Mayer found four propositions that characterized and described the type of information in the problem. These four basic propositions were: (1) the assignment proposition, (2) the relation proposition, (3) the question proposition, and (4) the relevant fact. The way in which the structure and format of information is formulated in the problem defines the generic type of word problem. Each of the above points will be expanded more fully in the body of this paper, as well as several other views of algebra word problems.

In terms of Mayer's (1982) problem types, we examined the propositional relation algebra word problem using Kintsch and Greeno (1985) template schema to describe elements in the text. We adapted the Kintsch and Greeno model to assign the
key contextual features to the propositional relation problems. Furthermore, the propositional relation problems were constructed so that they had different presentation and responding modes (i.e., pictorial, symbolic and verbal) in combination with the key contextual features of familiarity, imageability, and variable type (discrete and continuous type), which were the major key contextual features found in the literature.

Logical analysis revealed that when problem presentation and responding modes were crossed with the key context features of familiarity, imageability and variable type, a large set of problems were needed in the set to represent all combinations. Thus, for each "cross translation" from a given presentation to response mode representation there would be 24 items to validate. In addition to these items, a parallel set of 24 items would have to be constructed to have "fresh" items for students to do on a post-test, which would create practical problems in assessing these word problems for reliability and validity. Therefore, a domain referenced (Hively, Maxwell, Rabehl, Sension and Lundin, 1973) set of problems was constructed which are discussed below. These problems will be used to study several variables that may be important to understanding the nature of the difficulties students have when trying to solve a propositional relation type of algebra problems.
Introduction

The construction of word problems with key contextual features has been addressed in several research studies (e.g., Caldwell, 1984 and Caldwell & Goldin, 1984). These studies have defined the key contextual features (KCF's) as being unrelated to elements of the algebra problem that make up the problem's mathematical structure. We have attempted to vary the KCFs in algebra problems using Kintsch and Greeno (1985) framework for describing template schemas for analyzing a "story line." The central point of this systematic variation of KCFs is that KCFs influence various aspects of the cognitive processing of algebra word problems. Stated in another way, KCF's (i.e., stimulus content) can control the cognitive demands that effect the solution of the problem.

The use of algebra word problems by mathematics educators to study student problem representation and translation behaviors has been extensive (e.g., Simon & Paige, 1966; Hinsley et al., 1977; Rosnick & Clement, 1980 and Clement, 1982). These research studies have viewed student representations and solutions of the problem posed as a product of the verbal structure of the algebra problem depicted in a relational proposition of some kind (Mayer, 1982). None of these studies, however, have reported the rationale used for constructing the algebra problems and how KCFs are varied in the problems.

Typically, research studies on problem-solving in
mathematics education have investigated the end result, or the solution of a problem isolated from the stimulus. The exact nature of the stimulus (i.e., the problem) in terms of it's key contextual features tends to be ignored by those doing research in this area and assessed to be a factor of little consequence or influence. We, obviously, are contending that this "unimportance" is not the case.

In most cases, researchers simply make assumptions about the content and structure of algebra word problems. The assumptions about the algebra problems are based on introspection that have usually ended with the problem itself being replicated with small variations in its content and structure and no systematic use of KCFs have been established as being important by works in the field. We, therefore, are going to present a schema for the construction of a set of algebra problems with key contextual features that are systematically varied. These key contextual features are: (1) familiarity, (familiar and unfamiliar); (2) imageability, (readily imageable and not readily imageable) and (3) variable type (discrete and continuous). These KCFs are varied in conjunction with three presentation and responding modes. These modes are: pictorial, symbolic, and verbal. Consequently, six types or modes of "cross translation" are possible; for example, translation from a presentation mode (such as verbal) to a responding mode (such as pictorial). The variations in KCFs and presentation and response modes are given in detail in Table 1 and Table 2.
**Rationale**

According to Mayer (1982), information given in algebra problems can be represented as a list of propositions. Mayer identifies four types of algebra word problems. One type of the four algebra word problems type, is the propositional relation problem. It entails describing the relationship between two covarying variables. This type of problem has been used in a verbal format to study student responses and errors (e.g., Paige & Simon, 1966; Rosnirk & Clement 1980; Clement, 1982; Mestre, Gerace and Lochhead 1982; Wollman 1983; Gerlach 1986; Mestre & Gerace 1986; Niaz 1989 and Sims-Knight & Kaput 1983a, 1983b).

A classic example of a propositional relation problem is the "professors and students" presented in verbal form. It goes as follows:

"Write an equation, using the variables S and P to represent the following statement: There are six times as many students as professors at this university. Use S for the number of students and P for the number of professors."

None of the investigators listed above have attempted to present this "standard" propositional relation problem in the different modes of representation that were proposed by Brunner (1964); namely, in the pictorial and symbolic formats of both presentation and responding modes or formats. Khoury and Behr (1982) state that the translation of relationships from one mode (i.e., pictorial, symbolic and verbal) to another should enhance meaningful learning, retention and transfer of mathematical ideas. However,
Brunner's original idea that cognitive development is related to information presentation and responding modes is equally important and in need of study, particularly in relation to algebra problems.

Research studies also have not systematically varied the key context features of problems (i.e., the non-mathematical information embodied in a problem). Sims-Knight and Kaput (1983a and 1983b), Caldwell (1977) and Horwitz (1980), operating from an information processing model, underlined the importance and use of the context dimension in the problem to understand student performance and comprehension of the translation task. However, almost all the reviewed literature characterized the context dimension only in the verbally presented problem. We, therefore have attempted to broaden the range of problems of the propositional relation type in existence and available for use with some limitation to the domain.

**Types of Presentations and Translations**

Few studies have investigated student problem-solving performance on the translation of a problem from one mode of representation to another. One unique study by Clarkson (1978) focused on the general performance of students on algebra problems represented in pictorial, symbolic, and verbal forms. Her algebra problems, and the modes of translations in them, however, were confounded. The problem's structure along the presentation mode was not homogeneous. Furthermore, Clarkson did not consider the
effects of problem features such as the syntax and context. Nevertheless, she was the first to define types of translations with respect to the cognitive processes required.

Clarkson called one type of translation an "active" translation. This type of translation includes verbal to pictorial, verbal to symbolic, and pictorial to symbolic problem translations. That had to been done actively by the student. A problem where the student selected the correct solution from multiple solutions provided in different forms was called a "passive" translation problem by Clarkson. Using Clarkson's idea (and problem feature), we have incorporated two types of translations in our algebra word problems; namely generative (or active translations) and passive translations. Generative translations are when a given problem is in one representative form and students are required to give (generate) an answer in an alternative form. Passive translations are when given one form, students select the correct answer from several other possible forms. This notion of generative and passive responding can be directly tied to both cognitive and psychometric theory. Table 1 is adapted from Clarkson study (p.6) which shows the translation of a problem from its original form to its resulting one.

**Key Context Features**

Key contextual features in algebra problems are defined as information that is embedded in the problem that might be necessary or unnecessary for the solution of the target
problem, but which are independent from the problem's structure and syntax. Contextual features in problems are sometimes implicit and elusive. Also, they can provide information that may be nonsensical or inconsistent with reality. For example, the following problem was presented by Paige and Simon (1966): "the number of quarters a man has is seven times the number of dimes--how many has he of each coin?" The information which the latter problem presents is contradictory and thus not possible. Context, then is a term that describes the non-mathematical meanings or content present in the problem statement. Context information, however, may help induce meaning to the mathematical content (Kulm, 1984, p. 17).

Contextual features conveyed by words and grammatical structures in a problem statement directly relate to the depth of encoding that students undertake to relate elements in the problem. Paige and Simon (1966) showed that when auxiliary information (i.e., additional information that is not used for eliciting strategies for a solution) was embedded in a problem, the information was not used as part of the translation process. They gave this example: "There are seven times as many quarters as dimes." Students who were interviewed in the study failed to recognize that extra information was needed to solve the problem.

In a preliminary study, we administered several algebra problems with auxiliary information to a group of college trigonometry students which contained one item adapted from
Simon and Paige's exploratory tests. All the students (n=27) who had taken algebra failed to disregard the auxiliary information (i.e., objects or noun referents) in developing a correct solution to the problem. Auxiliary information here was information that was redundant and irrelevant to the solution of the problem.

The elaboration on the referents of problems, such as the "quarters and dimes" problem given in Paige and Simon, require knowledge of what these quantities represent as indivisible nominal entities in their contextual domain. Following up on Paige and Simon's work, Kaput (1987b) states that "the elaboration [of problems] is done by using the features of the reference field of the symbol system rather than using its symbol scheme syntax (p.177)." This means that by "seiving" out redundant information (i.e., contextual attributes connected to noun referents and objects), the problem may become much easier to handle as its elements are more clearly organized and understood.

In the literature, KCFs are found associated with arithmetic (one-step) and algebra (two-step) word problems. The KCFs tend to be contrasts such as concrete-abstract, real world-fictitious, and familiar-unfamiliar. These KCFs, however, are defined in vague, inconsistent, and contradictory ways in the literature. In addition, those studies found in the literature which consider KCFs associated with the everyday reality (e.g., Washborne & Morpett, 1928; Houtz 1973, Caldwell, 1977 and Quintero, 1980)
used contexts based on children's experiences with concrete materials (regularity of the stimulus) moving on a continuum that ended in abstract or hypothetical modes. Thus, none of these studies have used KCFs as they relate to the pedagogical events in students' common instructional experiences (e.g., the use of discrete and continuous quantities).

As we are operating from an information processing perspective, we hypothesize that concrete attributes in the individual's cognitive structure should either entice the solver to process the problem, or arouse well anchored and familiar ideas in the individual's cognitive structure which will aid in the production of a direct representation of the problem.

We have identified three important KCFs from the experimental literature. These 3 KCFs are: (1) familiarity, (familiar and unfamiliar), (2) imageability, (readily imageable and not readily imageable) and (3) variable type (discrete and continuous quantities). Each of these KCFs are defined in Appendix A. These KCFs are logically crossed with a presentation and response mode (pictorial, symbolic, and verbal) of the propositional relation problems. It should be noted, however, that some combinations of features are difficult to construct actual word problems for, and that some combinations of features are more important than others theoretically, relative to obtaining experimental data. More will be said on each of these points below.
The Use of Key Context Features

Our focus on two key context features (i.e., familiarity and imageability) derives from their use in psychological literature on memory effects. These features have been used to study the recall of individuals on complex verbal material. Normative ratings of KCFs have been utilized for nouns (Stratton, Jacobus & Brinley, 1975; Rubin, 1980) transitive verbs (Klee and Legge, 1976) and adjectives (Berrian, Metzler, Knoll and Clark-Meyers 1979) by judges. Further, key contextual features (i.e., familiarity and imageability) have been studied with more complex structures such as proverbs and sayings (see Cunningham, Ridley & Campbell, 1987; and Higbee & Millard, 1983). None of these studies, however, have employed or attempted to employ any explicit theory to explain or predict the effects of KCFs on the process of solving the problem posed.

Familiarity as a key contextual feature has been used in studying algebra and arithmetic problems since 1926 (see Washbrone & Osborne, 1926; Washbrone & Morpett, 1928; Brownell & Stretch, 1931 cited in Webb 1984; Sutherland 1942 and Lyda & Franzen 1945.) These studies have shown that students more readily solve one and two-step arithmetic and algebra problems that have familiar referents than unfamiliar referents.

Imageability as a key contextual feature in algebra problems has not been investigated extensively as it relates to the problem-solving process. Imagery is considered to be
a mental picture that is formed in the mind prior to identification of a new one. The mental picture analogy describes a "rehash" of previous stored sensory patterns (Kosslyn and Pomerantz, 1977). Bramhall (1939) and more recently, Sims-Knight and Kaput (1983a) studied the effects of three types of imagery on solving relational proposition problems. Research in this area, however, has confused spatial ability with the notion of imagery as it relates to mathematics problem-solving. Clements (1981) stated that these two components of cognition (i.e., spatial ability and imagery) have qualitatively different processing strategies. Furthermore, as with the case of familiarity, those studies which have used imageability as a key context feature have not usually defined this feature with respect to some theory, and most of the studies have been rambling and inconsistent with one another.

The variable type KCF has not been investigated extensively in the literature at all. Horwitz (1980) used discrete and continuous variable types to indicate levels of visualization in problems. In her study, she defined discrete quantities as those which can be counted and continuous quantities are those that are measured. In some instances, problems dealing with intensive quantities, such as velocity, are not measurable, though their units are measured. Continuous quantities are considered to be problem elements that cannot be imagined (e.g., the weight of a box), as opposed to discrete quantity (i.e., "two boxes" or "two
oranges") which can be easily imagined. Further, continuous quantities are variables that can take on numerous values, such as units of currency, whereas apples and horses are cardinal values that have no numerical value.

Clements (1981) found that higher error rates that occurred on problems with continuous quantities. As Horwitz (1980) noted, discrete variables are easily visualized, as they are easier to encode. Discrete elements are usually real objects and events that are countable into wholes. Propositions that have continuous events are not easy to imagine and this characteristic is considered to be one of the major reasons that such propositions cause performance difficulties.

**Propositional Representation of Knowledge**

Given the literature in this area, we attempted to develop a set of algebra word problems which have a propositional relational structure and the KCFs of familiarity, imageability, and variable type. These three KCFs are systematically varied along three different representational dimensions (verbal, pictorial and symbolic). The interpretation of students' performances on these problems is based on a theory proposed by Simon (1972), Anderson & Bower (1973) Pylyshyn (1973) and Norman (1981). This theory states that any knowledge can be represented by a set of propositions. Further, there are macro and micro propositions in this view. The macro propositional elements are linked by a stipulated (defined and specific)
relationship. Within this macro proposition we find micro-propositions which form a logical set of objects, quantities, specifications (qualifiers, determinants of roles) and relations. This theory further asserts that verbal, pictorial or symbolic representations as well as any form of representation (i.e., mixed models) can be transformed into propositions.

A perfect illustration of this point is a computer graphics program where the picture or image seen is produced by a set of computer statements, each of which is a proposition. A student solving a pictorial, symbolic or verbal problem with salient imageable attributes he/she searches for the proposition which represent the imageable or familiar component. This proposition is then transformed into verbal information that is imageable. In the case were the abstract representation system, such as a symbolic representation, does not contain a proposition relevant to a required piece of information, a person may deduce this proposition from those which are available. This type of situation and behavior would occur in and is indicative of a not readily imageable context as defined in this paper. Both successful performances and failures on word problems with specific KCF, therefore, can be modeled and analyzed from this propositional representation framework and point of view.

The Problems

We initially constructed 16 algebra word problems.
These problems represented all translation combinations; namely, verbal-pictorial (VP), verbal-symbolic (VS), pictorial-verbal (PV), pictorial-symbolic (PS), symbolic-verbal (SV) and symbolic-pictorial (SP). In terms of the other key contextual features, the PV, SV, and SP combinations are all of passive translations and their converses (the VP, VS and PS combinations) are generative translations.

Logical analysis revealed that problems were extremely difficult to construct for each presentation and response format that had all values of familiarity, imageability and variable type. The fully crossed "logical design", therefore, reduced to triads (see Table 2, for details). The key contextual feature triads of (1) familiar-readily imageable-discrete, (2) familiar-not readily imageable-continuous, (3) unfamiliar-readily imageable-discrete and (4) unfamiliar-not readily imageable-continuous could only be constructed for the verbal to symbolic and verbal to pictorial representation. The unfamiliar-readily imageable-discrete and unfamiliar-not readily imageable-continuous feature triads, however, could be constructed for all the presentation and responses formats (see Table 2, for details).

Four problems of the VS and four of the VP type were constructed. These two type of problems required generative translation. Four of the VS type of problems had the FI, FU, UI and UU problems attributes, and four of the VP problems
had the corresponding FI, FU, UI and UU attributes. Hence, two problems for each of the SV, PV, PS and SP were constructed having the UI and the second a UU assignment (see Table 2). Based on the translation type, the SV, PV and SP were passive translation problems and the PS, VP and VS were generative translations (see Table 3, in conjunction with table 4).

Assigning Key Context Features to Problems

An adaptation of the story problem schema developed by Kintsch and Greeno (1985) was employed to identify the KCFs and assign them problems. According to Kintsch and Greeno (1985), a student's set schema is a major factor or component in the problem-solving process as it deals directly with the structure of the algebra problem; i.e., the general form of the "story line." A student's set schema has four attributes according to Kintsch and Greeno (1985). These attributes are: object, quantity, specification and role. The object attribute is the set of all the noun referents; i.e., the type of elements or noun objects in the problem. If an object is not common (i.e., unfamiliar) then the object frame makes a schema that relies on resources during the process of comprehension. The quantity attribute is the cardinality of the set; e.g., "6 professors" or "10 students." The quantity is considered to be a qualifier. In this view, the specification attribute is a problem component that holds information connected to the quantity; e.g., owner, location and time case. The role attribute explains the relation
between the object and its modifier; e.g., which is the large set or small set in "6 professors" and "10 students."

Because Kintsch and Greeno (1985) problem-solving model is specific to one-step and two-step arithmetic problems, this schema was adapted and expanded so that it could apply to the propositional relational statement and facilitate the assignment of KCFs attributes to problems. This adaptation and expansion was done follows. The elements of a "story line" for the propositional relation problem are Noun referents, Qualifiers, Quantities and Relationships. The noun referents refers to the objects in the problem statement; i.e., apples, oranges, professors and so on. Qualifiers function as determiners as well adjectival modifiers; e.g., "stupid professors or 2.2 students." Other examples are "speed of a car" or "length increase of a box" which function as determiners of nouns. Thus, single noun referents cannot function independently as determiners. Quantities are the adjective modifiers, they express the cardinality of the objects or the number. For example, 2.2 professors or 6 students, 2.2 and 6 are the quantities. Relationships connect the quantities of the noun referents into a proportion e.g., "there are 20 students for one professor."

A verbal direction must logically prefaces each problem (e.g., write an equation, using the variables S and P). This direction (which is a proposition) is verbally stated. In all of the problems, the macro-propositional direction
precedes a symbolic, pictorial or verbal representation of the problem. The propositional direction states that the answer to the problem necessitate a pictorial, symbolic or verbal translation.

In processing the problem, the problem solver must relate the propositional statement or directions to the symbolic, pictorial or verbal representation of the problem. For example, in the symbolic presentation "3X=4Y," X and Y will be specifically denoted in the problem by a qualifier or/and noun referent (e.g., X stands for the number of apples and Y stands for the number of oranges). Similarly, pictorial presentations will depict the relationship between the two variables, and will also be specifically denoted by a noun referent and qualifiers in the problem.

Assignment Rules

There are three specific rules for determining if a problem feature is familiar or unfamiliar. First, if the noun referents in the algebra problems are familiar but the relationship is unfamiliar (e.g., 4 wheels for every 4 cars) then the problem feature is considered to be unfamiliar. Second, if the noun referents are unfamiliar (e.g., "donks" and "bonks"), then the quantitative relationship in the algebra problem feature is considered to be unfamiliar and hence an unfamiliar problem feature. Lastly, if the noun referents are familiar and the relationship is familiar (e.g., four wheels for each car), then the problem feature is considered to be familiar.
In a propositional statement such as "the number of 15 cars that have 15 wheels," the noun referents (i.e., cars and wheels) may be familiar to the problem solver but the ratio between the number of wheels to the number of cars is atypical. Such a proposition would be classified as unfamiliar, by our classification rules.

The assessments of the familiar or unfamiliar features should be done over the whole problem domain; namely, in terms of the Quantitative relationship rather than just a single assignment in the problem. The assessments of the imageable or not readily imageable features should be done based on the noun referents. However, it should be noted that noun referents do not function independently of their qualifiers, and in some instances problems may have qualifiers which decide the imageability of the problem.

Given, this point there is one specific rule to determine if each algebra problem is readily imageable or not readily imageable. If the qualifiers are determinant of the noun referents, then judging the problem as readily imageable or not readily imageable should be done over all features of the problem, using the qualifiers and noun referents (e.g., "speed of the car" versus "number of cars"). In the later example, speed plays a role as a qualifier which makes the first proposition harder to imagine as opposed to the second.

Classifying the type of variable (i.e., discrete or continuous) in a problem is done from the relationship and
qualifiers. There are two views of any variable. The first view is a relatively static one. It is associated with algebraic symbols. How these symbols are connected to their qualifiers determines the type of variable or quantity.

The second view of a variable has a more dynamic character where one variable changes respective to another. The dynamic change can be represented by a graph. The x-axis and y-axis corresponds to the qualifiers which suggests the appropriate unit which determines the form of variable. In both modes, the variable may be either discrete or continuous depending on the qualifiers in the propositions. Often students tend to confuse variable type because of the format.

Most relations encountered in mathematics language are represented by a symbol-variable type. The variable symbol is used to denote the unknown, specifically the continuous or discrete quantities. In the context, which discrete and continuous quantities are directly linked to the domain of the problem determine variable type; i.e., it is the relation that is either discrete or continuous (Leinhardt, Zaslavsky and Stein, 1990).

An example here should facilitate understanding of how to apply this schema and how KCFs are assigned to problems. Given the following problem below:

The problem below relates the number of beads and the number of marbles in a container. B stands for the number of beads and M stands for the number of marbles. Choose one answer:

\[ 3B = 6M \]

a) There are twice as many marbles as there are beads.
b) There are three times the number of marbles as there are beads.
c) There are twice as many beads as there are marbles.

In the above problem, the beads and marbles in the statement are quantities. They are usually considered to be discrete but they could be continuous in a given context. The beads and marbles are either imageable or not readily imageable. For the vast majority of people doing this problem the beads and marbles would be imageable. The symbolic relation given by "3B=6M" is a presentation mode that depicts the relation and is used to determine if the algebra problem is familiar or unfamiliar. In this case, the situation is unfamiliar because by definition finding a specific number of beads related to a specific number of marbles is atypical. Though the problem is not presented in a spatial form, one can ascertain that the imageability of the propositional representation are readily imageable. Beads are objects which have certain characteristics like roundness. Beads are vivid and hence easily imaged. The variable involved is discrete because the number of marbles and number of beads cannot be decomposed into fractional parts.

Imageability should not "feel" pictorial or that anything is occurring in the mind that is pictorial. For example, some problems are pictorially presented. These problems should be considered to be either readily imageable or not readily imageable. Readily imageable then is, then, a process which is concerned with an immediate mental picture of the objects involved without a generative construction.
that produces a diagrammatic representation of a concept.

Familiarity deals with the high frequency of exposure to some structural or imageable form prior to one's new experience with that form. The relation of the number of beads to the number of marbles is and should be considered unfamiliar.

Thus, some propositional statements differ in their ability to arouse visual images of objects or events. Some propositions elicit a mental picture which is very easily imageable where others that are more abstract or unfamiliar are very hard to image. More detailed definitions of all these features and other examples of applying these features assignment are given in Appendix A. These directions were used for training judges to rate our algebra word problems in terms of their key features. The rating of six judges agreed with our "claimed" features for the 16 problems we constructed to date 96% of the time (see Nasser and Carifio, 1993a). The construct validity of our conceptualization and rules, therefore, are quite good if not excellent.

Closing Remarks

A rationale and theory has been presented for the construction of algebra problem sets specific to a structural domain. The problems we have constructed so far may be used to study the most important dimensions of student qualitative and quantitative reasoning (see Nasser and Carifio, 1993b). To date we have constructed sixteen algebra word problems with a minimum of another 8 to be constructed. In terms of
the 16 problems we have constructed, the KCFs are nested in terms of presentation format and type of translation (see Table 5). Four problems are of the verbal to symbolic and four are of the verbal to pictorial cross translation type. There is two problem corresponding to four key context feature triads. These triads are: (1) familiar-readily imageable-discrete, (2) familiar-not readily imageable-continuous, (3) unfamiliar-readily imageable-discrete and (4) unfamiliar-not readily imageable-continuous. The symbolic and pictorial problems were limited to the triads of unfamiliar-readily imageable-discrete quantities and unfamiliar-not readily imageable-continuous quantities (see Table 2 and Table 3).

The 16 algebra word problems we have constructed to date do not comprehensively complete the propositional relation problem domain with all the qualities, although different problems do systematically vary the same concepts key features (i.e., problems are presented in pictorial, symbolic and verbal formats which contain similar key contextual features). The 8 word problems needed to complete the domain (see Table 1) are in the process of being constructed and validated. Not all the problems shown on the matrix were constructed in this study. Only 16 problems have been constructed and validated so far. As indicated by Hively, Maxwell, Rabehl, Sension and Lundin (1973), simple concepts have so “many potential representative” behaviors that it is impossible to specify them all. Hence, limits to the problem
sets and domain must be established. The structured sets of problems and the domain, with all of its properties (i.e., generative, passive, pictorial, symbolic and verbal) and key contextual features could not be equally and comprehensively developed because such a domain would encompass a very large number of items. It would be incomprehensibly exhaustive to both develop and generalize about such a problem sets. Our reduction of this complex logical domain to a parsimonious subset that represents the most important triads, made the domain tractable for research, theory building and program evaluation activities. However, within the domain of the propositional relation problem, researchers may refer to Table 2 and generate other problems constituting this domain as they are needed.

The validity and reliability of the 16 algebra word problems constructed to date as done in two phases. In phase one, four judges reviewed the 16 problems constructed in terms of the feature rules outlined here and the template for the story-line schema of a propositional relation structure. Following this first phase, six judges rated the 16 problems in random order in terms of the KCF of the problems outlined here. As previously stated, the ratings of the six judges agreed with our claimed features of the 16 problems, 96% of the time. The construct validity of our conceptualization and operationalized rules is, therefore, quite good and valid word problems, which have various combinations of the key contextual features of word problems outlined in this paper,
can be readily and easily constructed by researchers and curriculum writers.

As can be seen from the work we have done on algebra word problems, the research literature, logical analysis, and learning theory can be used to model a complex phenomenon (or behavior) and its domain in a systematic and rigorous manner. Once conceptualized, operational definitions and rules may be developed which facilitate the generation of problems (or items) which systematically vary the key features of the defined domain. Using a panel of judges, the items may be validated in terms of their claimed characteristics. Once validated, the characteristics of each problem (or item) as a stimulus (or criterion behavior) is known with some reasonable precision, which means that "cause and effect" and the results of experiments, evaluations, and curricula may be better assessed. Currently, the latter problem is a major problem in research and evaluation in mathematics education as well as other educational areas. This problem is only going to become worse with the "rapid jump" to "authentic assessment" that is occurring in education currently. Our work, therefore, is a model for others to employ.

Now that we have developed and validated a set of algebra word problems to use (see Nasser and Carifio, 1993a), we are proceeding with efforts to conduct studies on how various individual differences variables such as level of cognitive development, field dependence, verbal, nonverbal and quantitative abilities influence performance on algebra
word problems with different key contextual features (see Nasser & Carifio, 1993b). These studies will help us develop and validate an information processing model of algebra word problem solving behavior which will have direct instructional implications. The point here, however, is that if we had not employed the model outlined here to develop and validate a domain-reference set of algebra word problems which systematically varying the key contextual features we identified from the research literature and logical analysis, we would not have to conduct these latter studies. Conceptualizing, theorizing and reviewing the existing literature, therefore, is the first and not the last step in the research process.
TABLE 1:
RESPONSE MODE FORMATS BY THE FORM OF RESPONSE

<table>
<thead>
<tr>
<th>Presentation Mode</th>
<th>Verbal</th>
<th>Pictorial</th>
<th>Symbolic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal</td>
<td>------</td>
<td>Generative</td>
<td>Generative</td>
</tr>
<tr>
<td>Pictorial</td>
<td>Passive</td>
<td>------</td>
<td>Generative</td>
</tr>
<tr>
<td>Symbolic</td>
<td>Passive</td>
<td>Passive</td>
<td>------</td>
</tr>
</tbody>
</table>

TABLE 2:
A DESCRIPTIVE AND CONCEPTUAL CHARACTERIZATION OF THE DOMAIN OF ALGEBRA WORD PROBLEMS

<table>
<thead>
<tr>
<th>Mode of Representation and Cross Translation</th>
<th>Key Contextual Features Triads</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FI/D</td>
</tr>
<tr>
<td>Verbal to Symbolic</td>
<td>1</td>
</tr>
<tr>
<td>Symbolic to Verbal</td>
<td>2</td>
</tr>
<tr>
<td>Pictorial to Symbolic</td>
<td>2</td>
</tr>
<tr>
<td>Symbolic to Pictorial</td>
<td>2</td>
</tr>
<tr>
<td>Verbal to Pictorial</td>
<td>1</td>
</tr>
<tr>
<td>Pictorial to Verbal</td>
<td>2</td>
</tr>
</tbody>
</table>

FI/D= familiar-readily imageable-discrete
UI/D= unfamiliar-readily imageable-discrete
FU/C= familiar-not readily imageable-continuous
UU/C= unfamiliar-not readily imageable-continuous
1= First set of word problems to be developed
2= Expansion set of word problems to be developed
<table>
<thead>
<tr>
<th>Presentation and Translation Type</th>
<th>FI</th>
<th>FU</th>
<th>UI</th>
<th>UU</th>
<th>Problem Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS-G</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>VS2</td>
</tr>
<tr>
<td>VS-G</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>VS3</td>
</tr>
<tr>
<td>VS-G</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>VS1</td>
</tr>
<tr>
<td>VS-G</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>VS4</td>
</tr>
<tr>
<td>VP-G</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>VP2</td>
</tr>
<tr>
<td>VP-G</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>VP4</td>
</tr>
<tr>
<td>VP-G</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>VP1</td>
</tr>
<tr>
<td>VP-G</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>VP3</td>
</tr>
<tr>
<td>PS-G</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>PS1</td>
</tr>
<tr>
<td>PS-G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PS2</td>
</tr>
<tr>
<td>PV-P</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>PV1</td>
</tr>
<tr>
<td>PV-P</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>PV2</td>
</tr>
<tr>
<td>SP-P</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>SP1</td>
</tr>
<tr>
<td>SP-P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SP2</td>
</tr>
<tr>
<td>SV-P</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>SV1</td>
</tr>
<tr>
<td>SV-P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SV2</td>
</tr>
</tbody>
</table>

P= Passive Translation  
G= Generative Translation  
SV= Symbolic To Verbal  
PS= Pictorial to Symbolic  
SP= Symbolic to Pictorial  
VP= Verbal to Pictorial  
PV= Pictorial to Verbal  
P= Passive Translations  
G= Generative Translations
<table>
<thead>
<tr>
<th>Item Code</th>
<th>Labels</th>
<th>Problem Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS1</td>
<td>V-S/G/UI/D</td>
<td>Verbal-Symbolic/Generative/unfamiliar-readily imageable-discrete quantity</td>
</tr>
<tr>
<td>VS4</td>
<td>V-S/G/UU/C</td>
<td>Verbal-Symbolic/Generative/unfamiliar-not readily imageable-continuous quantity</td>
</tr>
<tr>
<td>SV1</td>
<td>S-V/P/UI/D</td>
<td>Symbolic-Verbal/Passive/unfamiliar-readily imageable-discrete quantity</td>
</tr>
<tr>
<td>SV2</td>
<td>S-V/P/UU/C</td>
<td>Symbolic-Verbal/Passive/unfamiliar-readily imageable-discrete quantity</td>
</tr>
<tr>
<td>PS1</td>
<td>P-S/G/UI/D</td>
<td>Pictorial-Symbolic/Generative/unfamiliar-readily imageable-discrete quantity</td>
</tr>
<tr>
<td>PS2</td>
<td>P-S/G/UU/C</td>
<td>Pictorial-Symbolic/Generative/unfamiliar-not readily imageable-continuous quantity</td>
</tr>
<tr>
<td>SP1</td>
<td>S-P/P/UI/D</td>
<td>Symbolic-Pictorial/Passive/unfamiliar-readily imageable-discrete quantity</td>
</tr>
<tr>
<td>SP2</td>
<td>S-P/P/UU/C</td>
<td>Symbolic-Pictorial/Passive/unfamiliar-not readily imageable-continuous quantity</td>
</tr>
<tr>
<td>VP1</td>
<td>V-P/G/UI/D</td>
<td>Verbal-Pictorial/Generative/unfamiliar-readily imageable-discrete quantity</td>
</tr>
<tr>
<td>VP3</td>
<td>V-P/G/UU/C</td>
<td>Verbal-Pictorial/Generative/unfamiliar-not readily imageable-discrete quantity</td>
</tr>
<tr>
<td>PV1</td>
<td>P-V/P/UI/D</td>
<td>Pictorial-Verbal/Passive/unfamiliar-readily imageable-discrete quantity</td>
</tr>
<tr>
<td>PV2</td>
<td>P-V/P/UU/C</td>
<td>Pictorial-Verbal/Passive/unfamiliar-not readily imageable-discrete quantity</td>
</tr>
<tr>
<td>VS2</td>
<td>V-S/G/FI/D</td>
<td>Verbal-Symbolic/Generative/familiar-readily imageable-discrete quantity</td>
</tr>
<tr>
<td>VS3</td>
<td>V-S/G/FU/C</td>
<td>Verbal-Symbolic/Generative/familiar-not readily imageable-continuous quantity</td>
</tr>
<tr>
<td>VP2</td>
<td>V-P/G/FI/D</td>
<td>Verbal-Pictorial/Generative/familiar-readily imageable-discrete quantity</td>
</tr>
<tr>
<td>VP1</td>
<td>V-P/G/FU/C</td>
<td>Verbal-Pictorial/Generative/familiar-not readily imageable-continuous quantity</td>
</tr>
</tbody>
</table>
TABLE 5:
NESTED QUALITIES OF THE ALGEBRA PROBLEM

<table>
<thead>
<tr>
<th>Translation:</th>
<th>Generative</th>
<th>Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presentation:</td>
<td>VP PS VS</td>
<td>SV PV SP</td>
</tr>
<tr>
<td>KCFs:</td>
<td>FI FU UI UU</td>
<td>UI UU UI UU</td>
</tr>
</tbody>
</table>
Reference


Rules and Definitions for Assigning Key Context Features to a Propositional Relation Problem

Attached are 16 algebra problems I am going to ask you to read and evaluate each algebra problem in terms of whether or not certain attributes or features are present or absent in each problem. There are 3 criteria for evaluating each word problem each of which will have 2 attributes or features.

These 6 features, therefore are familiar, unfamiliar, readily imageable or not readily imageable and discrete or continuous quantities.

After reading and evaluating each word problem according to the rules given below, you will check the attributes that you consider to be present in the problem on the scale, given beneath each problem:

<table>
<thead>
<tr>
<th>1. familiar _____</th>
<th>unfamiliar _____</th>
<th>not sure _____</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. readily _____</td>
<td>not readily _____</td>
<td>imageable _____</td>
</tr>
<tr>
<td>3. discrete _____</td>
<td>continuous _____</td>
<td>quantity _____</td>
</tr>
</tbody>
</table>

Definitions for each of the 6 features to be evaluated in each problem are given below. This is followed by the general rules for judging whether the features are present or absent in a given problem.

Definitions

Discrete Quantities: Are quantities which can be counted (e.g., dogs, cats, professors, students, etc.), that have no fractional parts. Discrete quantities in problem statements are considered to be elements that are physical, real and countable.

Continuous Quantities: Are quantities that may have fractional measurement or units. Speed is a continuous quantity with fractional units of distance and time which are also continuous quantities. Weight is a continuous quantity which has no decomposable parts but the unit itself can be expressed in fractional form e.g., 2.2 pounds or 3.3 kilos.

Familiar: Problem features (i.e., noun referents and
relationships of quantities) that are regular, have a high frequency of occurrence and are easily related to a previously known situation. Such a familiar situation is one where the experience has been seen on regular basis by a typical American adolescents and adults and is usually a part of h/his every day experience e.g., "one table usually has four legs," versus the unfamiliar situation of 20 tables with one leg or "one car with one wheel."

Unfamiliar: It follows from the above definition, where noun referents and quantitative relationships in the problems do not depict a familiar situation. The problem's quantitative relationships is irregular and atypical, for example "In a box there are twenty apples for every two oranges." Apples and oranges can be imaged but the situation where a certain quantity of apples is proportional to the quantity of oranges is probably not seen on daily basis, hence the relation between the number of oranges and apples is unfamilair.

Readily Imageable: Specific problem features (i.e., noun referents and qualifiers) that can be easily visualized e.g., "there are four wheels" or "twenty apples" Thus wheels and apples are single entities (i.e., noun referents) that have a depictable image and experienced on regular basis.

Not Readily Imageable: Specific problem features (i.e., noun referents) with attribute that are not readily imageable features e.g., use of noun words that do not exist in the English vocabulary or abstract representations denoted by words such as speed and acceleration that are difficult to visualize. One way to tell or identify features that are not readily imageable is that such features take a person longer time to process than readily imageable and familiar features because there must be an active process of reconstructing information, in order to make an image of the problem. For example, students may be familiar with the fact that the value of one dime is equivalent to two nickels. They can also image dimes and nickels, but the imageability of the value of dimes or two nickels is not immediate and not easily imaged. Similarly, pupil may be familiar with the speed of a sports car as it relates to the speed of a bus, but "speed" as a conceptual schema is not easily visualized.

General Rules

The rules for applying the features given above are:

1. In each algebra problem to be assessed there are four structural elements that will help you identify the features of the problem. These elements are Noun referents, Qualifiers, Quantities and Relationships. The noun referents refers to the objects in the problem statement i.e., apples, oranges, professors etc.. Qualifiers function as determiners
as well adjectival modifiers e.g., "stupid professors or 2.2 students." Other examples are "speed of a car" or "length increase of a box" which function as determiners of nouns, thus single noun referents cannot function independently as determiners. Quantities are the adjectival modifiers it holds the cardinality of the objects e.g., 2.2 professors or 6 students, 2.2 and 6 are the quantities. Relationships, connect the quantities of the none referents into a proportion of e.g., "there are 20 students for one professor."

2. There are three specific rules to judge if each problem feature is familiar or unfamiliar. First, if the noun referents in the algebra problems are familiar but the relationship is unfamiliar e.g., 4 wheels for every 4 cars then the problem feature should be considered unfamiliar. Second if the noun referents are unfamiliar e.g., "donks" and "bonks" then the quantitative relationship in the algebra problem feature should be considered unfamiliar hence, an unfamiliar problem feature. Last if the noun referents are familiar and the relationship is familiar e.g., four wheels for each car, then the problem feature should be considered as familiar.

In a propositional statement such as "the number of 15 cars that have 15 wheels," the noun referents (i.e., cars and wheels) may be familiar to the problem solver but the ratio between the number of wheels to the number of cars is atypical, hence unfamiliar. As mentioned, in this type of situation your assessments of the familiar or unfamiliar features should be done over the whole problem domain i.e., in terms of the Quantitative relationship, rather than just a single attribute in the problem.

3. The imageable or not readily imageable features of the assessments should be done on the noun referents. However, noun referents do not function independently of their qualifiers in some instances problems may have qualifiers which decides the imageability of the problem.

There is one specific rule to judge if each algebra problem is readily imageable or not readily imageable. If the qualifiers are determinant of the noun referents, your criteria to judging the problem as imageable or not readily imageable should be done over a domain of the problem. Using the qualifiers and noun referents e.g., "speed of the car" versus "number of cars." Here speed plays a role as a qualifier which makes the first proposition harder to visualize as opposed to the second.

4. The type of variable i.e., discrete or continuous quantities should appear from the relationship and qualifiers. There are two aspects to a variable, one interpretation is relatively a static one, it is associated
with algebraic symbols and how these symbols are connected to their qualifiers determines the type of variable or quantity. A second interpretation of a variable has a more dynamic sense where one variable changes respectively to another. The dynamic change can be represented by a graph, the x-axis and y-axis corresponds to the qualifiers which suggests the appropriate unit that determines the form of variable.

5. A relational proposition prefaces each problem. This proposition is verbally stated. In all of the problems the proposition precedes a symbolic, pictorial or verbal presentation. In making your judgements, you must relate the propositional statement with the symbolic, pictorial or verbal presentations. For instance, in the symbolic presentation "3X=4Y," X and Y will be denoted by a qualifier or/and noun referent (e.g., X stands for the number of apples and Y stands for the number of oranges). Similarly, pictorial presentations will depict the relation and will be denoted by a noun referent and qualifiers.

Example

Given the multiple choice problem, below:

The problem below, relates the number of beads and the number of marbles in a container. B stands for the number of beads and M stands for the number of marbles. Choose one answer:

3B=6M

a) There are twice as many marbles as there are beads.
b) There are three times the number of marbles as there are beads.
c) There are twice as many beads as there are marbles.

The beads and marbles in the statement are quantities (i.e., either discrete or continuous). The beads and marbles are either imageable or not readily imageable. The symbolic relation given by 3B=6M is a presentation that depicts the relation and used to judge if the algebra problem is familiar or unfamiliar. In this case the situation is unfamiliar, because the regularity of finding an amount of beads related to the number of marbles is not unique. Though the problem is not presented in a spatial form, you can ascertain that the visualization of the propositional representation are readily imageable because beads are objects which have certain characteristics like roundness, it could be given object resemblance, hence easily imaged. The variable involved is discrete because the "number of marbles and
number of beads cannot be decomposed into fractional parts.

Imageability, should not "feel" pictorial or that anything is occurring in the mind that is pictorial. For example, some problems are pictorially presented, these problems should be considered to be either, readily imageable or not readily imageable. Readily imageable is then a process which is concerned with an immediate mental picture of objects involved, without the generative construction that produces a diagrammatic representation of a concept.

Familiarity deals with the high frequency of exposure to some structural or imageable form prior to one's new experience with that form. The relation of the number of beads to the number of marbles is an and should be considered unfamiliar.

Thus, some propositional statements differ in their ability to arouse visual images of objects or events. Some propositions elicit a mental picture which is very easily imageable, where else others that are more abstract or unfamiliar are very hard to be imaged.