Abstract Title Page

Not included in page count.

Title: Designing and Redesigning a Framework for Assessing Students' Understanding of Foundational Fractions Concepts

Maria Mendiburo
Institute for Software Integrated Systems
Vanderbilt University

Laura Williams
Department of Special Education
Vanderbilt University

Robert Henson
Department of Educational Research Methodology
University of North Carolina Greensboro

Ted Hasselbring
Department of Special Education
Vanderbilt University

Background / Context:

The fact that research has shown that fractions are among the most difficult mathematical concepts for elementary school students to master (Behr, Harel, Post, & Lesh, 1992; Bezuk & Cramer, 1989; Moss & Case, 1999) provides a compelling motivation for research and innovation focused on improving the available assessment and instructional tools related to fractions concepts. To address students' deficiencies in learning fractions prior to beginning algebra, we began an IES-funded development project called *H*elping *A*t-risk students *L*earn *F*ractions (HALF[1]). The project aims to develop and pilot test a complete, technology-based fractions intervention program. The instructional model for the complete HALF intervention blends teacher-led instruction and technology-based instruction. As such, one major goal of the HALF project is to develop a computer system that can provide teachers with the data and instructional resources they need to design adapted fractions lessons that best meet the current learning needs of their students. A second major goal is to design a computer system that can provide differentiated computer-based instruction to students in the same class with different levels of prior knowledge of fractions. Both of these goals require using accurate measures to diagnose students' level of understanding of fractions prior to beginning the intervention, and they require being able to accurately track students' progress towards mastery during the intervention. However, we failed to identify any measures from outside sources that met the needs of our project, so we decided to create our own measures of foundational fractions concepts. To date, we've created 243 multiple-choice assessment items that measure multiple aspects of the content domain of fractions and have completed four rounds of pilot testing these items. In this paper, we report findings from our pilot testing that contradicted the content framework we developed at the beginning of the HALF project.

When we began the HALF project, we used previous theoretical and empirical literature about fractions (e.g. Behr, Harel, Post, & Lesh, 1992; Cramer, Behr, Post, & Lesh, 2009; Cramer, Post, & delMas, 2002; Kieren, 1976; 1980; 1988) along with mathematics content standards and benchmarks developed by several national organizations (e.g. Common Core State Standards Initiative, 2010; National Council of Teachers of Mathematics, 2000; 2006; National Mathematics Advisory Panel, 2008) to create the HALF Learning Objectives. These five learning objectives were intended to represent the foundational fractions concepts students need to understand before they can develop the ability to grasp more advanced rational number concepts (see Table 1). We used the same sources to develop a set of measurable skills and extension activities for each learning objective that we felt could be used to measure students' progress toward mastery of that objective (see Table 1). Both the objectives and the measurable skills and extension activities within each objective were designed to represent a linear progression between concepts and a logical order for teaching fractions using the HALF intervention. We used this framework to guide the development of the individual assessment items as well as the overall structure of a comprehensive diagnostic assessment that measures the breadth of students' knowledge of foundational fractions concepts.

Purpose / Objective / Research Question / Focus of Study:

This study focused on determining whether the theoretical framework developed in the

[1] The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A100110 to Vanderbilt University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

learning objectives is a good fit for the item data we have collected. We first attempted to fit a dichotomous Rasch model. The learning objectives were developed with a linear path in mind, such that the first objective should be the easiest content and the fifth should be the hardest, so we wanted to start with the simplest possible analysis modeling only the item difficulties. We then compared the unidimensional results to those of a multidimensional IRT model to determine if the item difficulties alone were sufficient to describe the data or if a multidimensional structure had a better fit.

Setting:

We chose a diverse group of school sites in order to ensure the results of our analyses generalized to the broadest range of settings possible. The sample currently includes participants drawn from intact classes at a private school for students with special learning needs, two public charter schools that serve primarily minority students from disadvantaged backgrounds, a private school without academic admissions requirements that has a one-to-one iPad program, a private school with academic admissions requirements, a public magnet school with academic admissions requirements, a traditional public school that serves primarily minority students from disadvantaged backgrounds, and two public schools with a heterogeneous student population. All of the schools in the sample are located in Middle Tennessee with the exception of the private school with the one-to-one iPad program, which is located in Eastern Tennessee.

Population / Participants / Subjects:

The sample includes 880 fifth- and sixth-grade students drawn from intact classes at 9 different schools. Roughly an equal number of male and female students participated, and the students represent a range of ability levels as measured by other standardized mathematics assessments. The sample includes students of multiple racial/ethnic backgrounds, students with identified learning disabilities, students who qualify for free or reduced-price lunch, and students classified as English-language learners. However, we are still in the process of collecting some of the demographic information for the sample and so at this time we cannot calculate the exact percentage of students in the sample that fall within each of these categories.

Intervention / Program / Practice:

We used an iterative process to develop and pilot test the 243 multiple-choice assessment items. We created the first pool of 60 items by searching the previously mentioned mathematics literature for examples of assessment items that aligned with one or more of the measurable skills or extension activities from the HALF Learning Objectives. We also searched the fractions content of three popular, commercial mathematics curricula (Burns, 2008; TERC, 2008; University of Chicago School Mathematics Project, 2004). We found high-quality examples for 15 measurable skills that aligned with the first four of five learning objectives (see Table 1). Based on these example items, we created four multiple-choice items for each measurable skill. The incorrect answer choices for each question reflected common misconceptions about fractions from the mathematics literature.

We used a similar process to create items for the next three rounds of pilot testing, but we also considered data from the pilot testing in the creation of new items. More specifically, after each round of item development and pilot testing, we examined the data to look for gaps in content coverage and contradictions between the item difficulty estimates and the framework we established in the HALF Learning Objectives. We then developed new items to address these gaps and contradictions. We also developed items that aligned with the fifth learning objective.

Research Design, Data Collection and Analysis:

We designed the platform-independent computer software that delivers our fractions assessment as an extensible web portal. We used a classroom set of iPad tablets and a closed network server operating from a laptop computer to collect data in intact classes of fifth- and sixth-grade students. Students' responses were submitted to the server via the closed network and saved in a database. After an assessment, standard database tools (*e.g.*, SQL scripts) were employed to extract and aggregate assessment information from the database into spreadsheets for further analysis.

We completed four rounds of pilot testing of the assessment questions in intact fifth- and sixth-grade mathematics classes. Multiple versions of the assessment were tested in each round, with each version comprised of 30 questions, some of which were previously tested anchor questions. Each student took only one version of the assessment, so the sample includes systematic missing data. All items were scored dichotomously, with an incorrect answer receiving 0 points and a correct answer receiving 1 point.

The data were first analyzed using a dichotomous Rasch model via Winsteps software (Linacre, 2012). A Rasch model describes an item properties using only a single item parameter that identifies the basic difficulty of the item, b_j. In addition examinee ability, θ_j, is defined using only a unidimensional measure of ability. The dichotomous Rasch model is defined as

$$\left(X_{ij} = 1, \theta_{\boldsymbol{i}}, b_j\right) = \frac{e^{\theta_i - b_j}}{1+e^{\theta_i - b_j}} \qquad (1)$$

where x_{ij} is the response of examinee i on item j, θ_i, is the ability estimate of examinee i, and b_j is the difficulty estimate of item j. The results of this analysis are described in the Findings section below.

After examining the results of the unidimensional Rasch, we attempted to fit the overall dimension structure of the five learning objectives using multidimensional IRT (MIRT; Reckase, 1997). The MIRT model is most generally defined as

$$p\left(X_{ij} = 1, \boldsymbol{\theta_i}, a_j, d_j\right) = \frac{e^{d_j + \sum_{s=1}^{S} a_{js}\theta_{is}}}{1 + e^{d_j + \sum_{s=1}^{S} a_{js}\theta_{is}}} \qquad (2)$$

where X_{ij} is the response of the i^{th} examinee of given ability vector estimate $\boldsymbol{\theta_i}$ on item $\mathbf{j}$, and $d_j$is the estimated difficulty parameter of item $\mathbf{j}$, which reflects the location of the item in reference to the ability scale. The items discrimination on dimension s is indicated by a_{js} and therefore a vector $\mathbf{a}_j$ is used to describe the discrimination of item j on all S dimensions. Mplus software (Muthén & Muthén, 2010) was used to estimate the overall fit of the multidimensional model and compare it to the relative fit of the unidimensional model via each model's AIC and BIC statistics.

Findings / Results:

The unidimensional Rasch model gave item difficulty estimates that did not match what we had expected. When we created the learning objectives, we hypothesized a linear progression of learning within and between the learning objectives, so that the difficulty parameter estimates would align with the progression of the objectives. That is, the items for Objective 1a would have the lowest (easiest) *b* estimates and the items for Objective 5f would have the highest (hardest) *b* estimates. Instead, we found that the actual item difficulty estimates were not well aligned with the objectives. While some difficulties were as expected (e.g. the items for Objective 1b, word names for common fractions, were among those with the lowest *b* estimates), other objectives' items were either harder than expected or had an unexpectedly large range of item difficulties. For example, an extension activity from Objective 2, comparing fractions using a number line, had items that were among the most difficult for students, and Objective 1a, partitioning objects into fractional parts, had item difficulties that were spread from very easy to relatively difficult. A strip chart displaying a sample of the item difficulties is shown in Figure 1.

Since the unidimensional Rasch model analysis yielded surprising results, we decided to test the dimensionality of the learning objectives as written using a 5-dimension MIRT model. The results showed that the multidimensional model (AIC = 27686.563) was a better fit for the data than the unidimensional model (AIC = 28196.956). However, the difference in the relative model fits is extremely small and the BIC, which gives more weight to the more parsimonious models, is lower for the unidimensional model (unidimensional BIC = 29401.817, multidimensional BIC = 30166.277). This indicates that there is effectively no difference between the unidimensional and multidimensional models, as the relative model fit statistics give a different answer depending on which is used. It is important to note that this is simply a comparison of two models, the unidimensional analysis of the item difficulties only and the multidimensional analysis of the structure outlined in the learning objectives. The AIC statistic does not indicate that the MIRT model used is the best possible multidimensional structure to fit these data, simply that at least one multidimensional model fits the data better than the unidimensional model. Given the misalignment between the observed and expected item difficulties in the Rasch model, we believe that there may be a different multidimensional structure that would fit the data better than the current objectives, one that could be achieved by revising the current learning objective framework.

Conclusions:

After considering the results of our pilot data analysis, we concluded that while the original HALF Learning Objectives may represent a logical order for teaching fractions to students who have little to no previous exposure to the content domain, it does not consider that the fifth- and sixth-grade students who are participating in HALF research have been learning fractions for three to four years but have different levels of understanding of the content domain. Within each of the original HALF Learning Objectives, there are concepts that fifth- and sixth-grade students find more challenging and concepts they find less challenging, so their understanding of the content domain will not necessarily align with a logical order for teaching the content for the first time. Therefore, we chose to create a new content framework for the HALF project (see Table 2). The new framework assumes that although some concepts within each strand and between strands are likely to develop in a linear manner, some students may develop fluency with certain higher concepts without developing fluency of the lower concepts because their understandings of some of the lower concepts are still forming. As they continue to learn the higher concepts, some lower concepts may be reinforced, and the student will continue to learn even though their path of understanding doesn't follow a linear progression. We plan to test this new framework in future research.

Appendices

Appendix A. References

Behr, M. J., Harel, G., Post, T.R., & Lesh, R. (1992). Rational number ratio and proportion. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 296-333). New York: Macmillan Publishing.

Bezuk, N., & Cramer, K. (1989). Teaching about fractions: What, when, and how? In P. Trafton (Ed.), *National Council of Teachers of Mathematics 1989 Yearbook: New directions for elementary school mathematics* (pp. 156-167). Reston, VA: National Council of Teachers of Mathematics.

Burns, M. K. (2008) *What is formative evaluation?* Minneapolis, MN: University of Minnesota.

Cramer, K., Behr, M. J., Post, T., & Lesh, R. (2009). *Rational Number Project: Initial fraction ideas*. Retrieved from http://www.cehd.umn.edu/rationalnumber project/rnp1-09.html

Cramer, K., Post, T., & delMas, R. (2002). Initial fraction learning by fourth- and fifth-grade students: A comparison of the effects of using commercial curricula with the effects of using the Rational Number Project curriculum. *Journal for Research in Mathematics Education*, *33*(2), 111-144. doi:10.2307/749646

Common Core State Standards Initiative (CCSSI). (2010). *Common core state standards for mathematics*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State Officers.

Kieren, T. (1976). *Number and measurement: Papers from a research workshop.* Retrieved from ERIC database (ED120027).

Kieren, T. (1980). The rational number construct: Its elements and mechanisms. In T. Kieren (Ed.), *Recent research on number learning* (pp. 125-149). Columbus, OH: ERIC/SMEAC.

Kieren, T. (1988) Personal knowledge of rational numbers: Its intuitive and formal development. In J. Hiebert and M. Behr (Eds.), *Research agenda for mathematics education: Number concepts and operations in the middle grades* (pp.162-181). Virginia: Lawrence Earlbaum Associates.

Lesh, R., Cramer, K., Doerr, H., Post, T., & Zawojewski, J. (2003). Model development sequences. In R. A. Lesh & H. Doerr (Eds.), *Beyond constructivism: A models and modeling perspective on mathematics, teaching, learning, and problem solving* (pp. 35-58). Mahwah, NJ: Lawrence Earlbaum Associates.

Linacre, J. M. (2012). Winsteps® Rasch measurement computer program [Computer software]. Beaverton, OR: Winsteps.com.

Muthén, L. K., & Muthén, B. O. (2010). Mplus (Version 6) [Computer software]. Los Angeles, CA: Muthén & Muthén.

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics Education, 30*, 122-147. doi:10.2307/749607

National Council of Teachers of Mathematics (2000). *Principles and standards of school mathematics.* Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence.* Reston, VA: National Council of Teachers of Mathematics.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Washington, DC: US Department of Education. Retrieved from www2.ed.gov/about/bdscomm/list/mathpanel/ report/final-report.pdf

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research, Copenhagen.

Reckase, M. D. (1997) The past and future of multidimensional item response theory. *Applied Psychological Measurement, 21*(1), 25-26. doi:10.1177/014662169 7211002

TERC. (2008). *Investigations in number, data, and space*. Cambridge, MA: Pearson.

University of Chicago School Mathematics Project. (2004). *Everyday mathematics*. Chicago, IL: McGraw-Hill Education.

Appendix B. Tables and Figures

Table 1.

Original HALF Learning Objectives

Objective #1: Students understand fractions as parts of unit wholes.
Measurable skills:
a) Partitioning whole objects into fractional parts.
b) Providing the word name for common fractions (e.g. one-fourth, two-sixths).
c) Choosing the correct fractional part given the whole.
d) Choosing the correct whole given a fractional part.
e) Constructing models of fractions using manipulatives.
f) Representing fractions using symbolic notation (e.g. 1/4, 2/6)
g) Extensions: word problems, partitioning a number line, representing common fractions on the number line.

Objective #2:Students can judge the size of fractions and generate equivalent forms of commonly used fractions.
Measurable skills:
a) Demonstrating the meaning of equivalence using models.
b) Recognizing and generating equivalent fractions.
c) Comparing fractions with like and unlike numerators and denominators, and recording the results of comparisons with symbols (<, >, =).
d) Using benchmark fractions (0, 1/2, 1) to compare fractions.
e) Ordering fractions.
f) Extensions: word problems, ordering fractions on the number line, demonstrating understanding of equivalence using the number line.

Objective #3: Students understand fractions as parts of a collection.
Measurable skills:
a) Partitioning groups of objects into fractional parts.
b) Using the context of a word problem to decide the appropriate model to use to represent a fraction.

Objective #4: Students can express whole numbers as fractions and have an understanding of fractions greater than one.
Measurable skills:
a) Modeling fractions equal to whole numbers.
b) Modeling fractions greater than one.
c) Representing fractions greater than one using mixed numbers and improper fraction notation.
d) Comparing mixed numbers and improper fractions.
e) Ordering groups of fractions that include fractions equal to whole numbers and fractions greater than one.
f) Extensions: word problems, representing mixed numbers and improper fractions on the number line, ordering mixed numbers and improper fractions on the number line.

Table 1 *continued*

Objective #5: Students can estimate computations involving fractions and can use models to solve computations.
Measurable skills:
a) Estimating the sum of two or more fractions.
b) Estimating the difference of two fractions.
c) Decompose a fraction into a sum of fractions with the same denominator in more than one way.
d) Using manipulatives to add and subtract fractions.
e) Using manipulatives to multiply a whole number by a fraction and a fraction by a whole number.
f) Using manipulatives to solve division problems that include a fraction in the quotient.
g) Extensions: word problems, modeling addition and subtraction of fractions on the number line, adding and subtracting mixed numbers with like denominators using models.

Table 2

New HALF Content Framework

Strand of Fractions Knowledge: *Identification*

Concept A: Represent proper fractions as parts of a whole.
Concept B: Represent proper fractions as parts of a set.
Concept C: Represent proper fractions as a number on the number line
Concept D: Represent mixed numbers.
Concept E: Represent improper fractions.

Strand of Fractions Knowledge: *Comparison and Equivalence*

Concept A: Recognize and generate equivalent proper fractions.
Concept B: Recognize and generate fractions that are equivalent to whole numbers.
Concept C: Convert improper fractions to mixed numbers and mixed numbers to improper fractions.
Concept D: Compare pairs of proper fractions with the same denominator or the same numerator
Concept E: Compare pairs of fractions with different numerators and denominators
Concept F: Compare pairs of mixed numbers
Concept G: Compare improper fractions and pairs of mixed numbers and improper fractions
Concept H: Compare groups of proper fractions, groups of mixed numbers, and groups of proper fractions and improper fractions
Concept I: Use common benchmarks (e.g. 0, ½, 1) to compare proper fractions and improper fractions
Concept J: Interpret and the relative size of fractions on a number line

Strand of Fractions Knowledge: *Addition and Subtraction*

Concept A: Estimate the sum and difference of proper fractions.
Concept B: Add unit fractions and decompose the sum of a fraction into unit fractions.
Concept C: Add and subtract proper fractions with the same denominator.
Concept D: Add and subtract proper fractions with different denominators.
Concept D: Add and subtract mixed numbers.
Concept F: Add and subtract improper fractions.

Table 2 *continued*

Strand of Fractions Knowledge: *Multiplication and Division*

Concept A: Multiply a whole number by a unit fraction and a unit fraction by a whole number.

Concept B: Multiply a whole number by a proper fraction and a proper fraction by a whole number.

Concept C: Multiply a proper fraction by a unit fraction.

Concept D: Multiply a proper fraction by a proper fraction.

Concept E: Predict the relative size of multiplication problems that include a whole number and a proper fraction, improper fraction, or mixed number.

Concept F: Divide unit fractions by whole numbers and whole numbers by unit fractions

Concept G: Divide a whole number by a whole number and interpret the remainder as a fraction.

Concept H: Predict the relative size of division problems that include a whole number and either a proper fraction, a mixed number, or an improper fraction.

We aim to assess deep understanding of these concepts, as operationalized by the Lesh (2003) Translation Model. By this definition, deep understanding of a mathematical concept requires experiences with different modes of representation and experience making connections between and within these modes of representation. In addition, the problem-solving activity required to translate a concept from one mode of representation to another reflects a dynamic view of instruction and learning. As such, students must demonstrate understanding of each concept of the above concepts using multiple modes of representation and must be able to translate between modes of representation as illustrated in Figure 2.

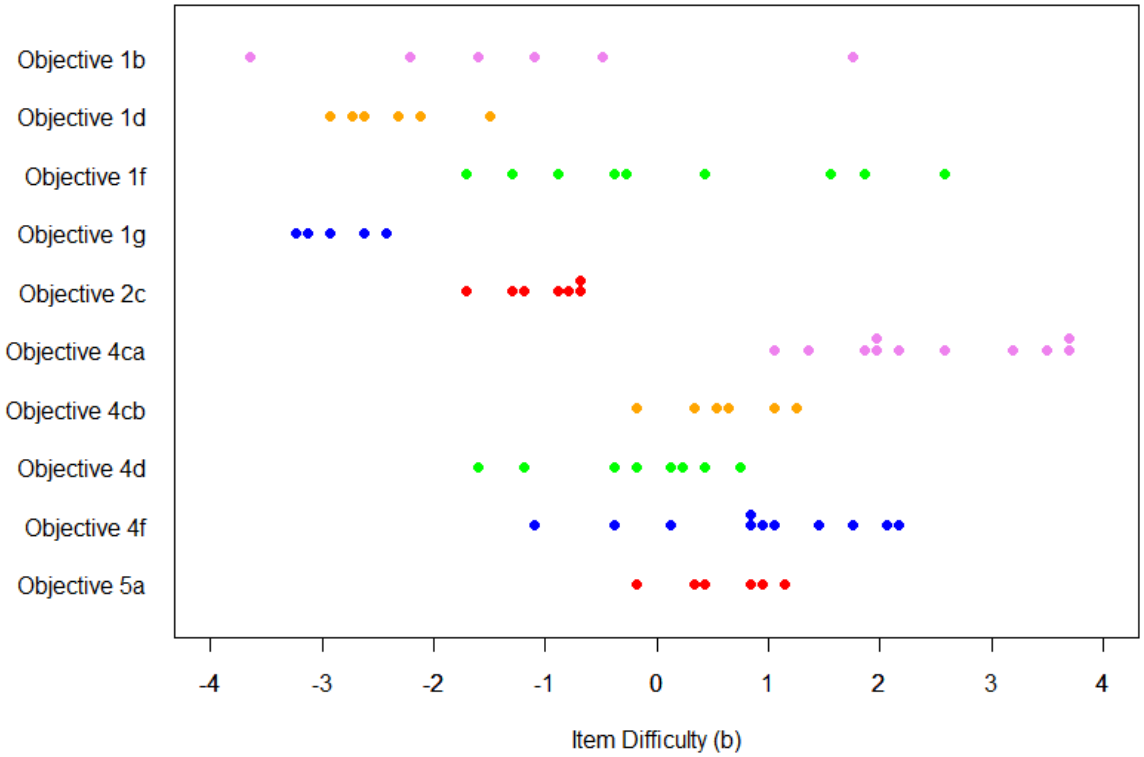

Figure 1. Strip chart showing a subsample of item difficulty estimates. According to the learning objectives (Table 1), the *b* estimates of the easiest (Objective 1) questions should have been clustered on the left and the *b* estimates of the hardest (Objective 5) questions clustered on the right.

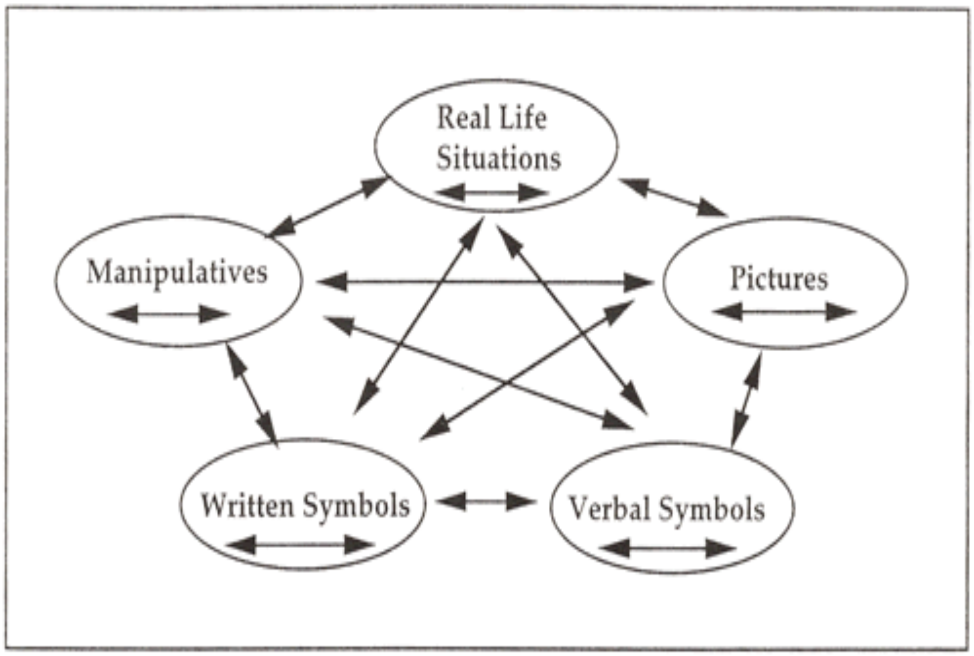

Figure 2. Lesh (2003) Translation Model depicting the different modes of interpretation between fractions concepts and connections between and within the modes.