## Abstract Title Page

**Title:** Half of One, 6/12 of Another: Understanding Relative Item Difficulties in a Fractions Assessment Under Development

**Authors and Affiliations:**

Laura K. Williams
Department of Special Education
Vanderbilt University

Maria Mendiburo
Institute for Software Integrated Systems
Vanderbilt University

Ted Hasselbring
Department of Special Education
Vanderbilt University

## Abstract


**Background / Context:**

A strong understanding of fractions is vital to later success in mathematics (National Mathematics Advisory Panel (NMAP), 2008). The fractions domain is a foundational content domain in learning algebra, which is in turn vital for success in advanced math and science courses (Gamoran & Hannigan, 2000; Kieren, 1976; United States Department of Education, 1997). However, research has consistently shown that fractions are one of the most difficult mathematical concepts for elementary school students to master (Bezuk & Cramer, 1989; Mullis, Dossey, Owen, & Phillips, 1991; NMAP, 2008). For example, an analysis of the 2000 NAEP Mathematics Assessment showed that only 41% of eighth graders were able to correctly order groups of 3 proper fractions, all of which were in reduced form (Kloosterman & Lester, 2004).

To attempt to remedy this deficit, we began an IES-funded development project called *H*elping *A*t-risk students *L*earn *F*ractions (HALF) in June of 2010. The purpose of the HALF project is to develop a technology-based learning environment that uses interactive visual representation to help students gain conceptual understanding of fractions. We intend to use this adaptive system to deliver differentiated instruction to students in the same class with different levels of prior knowledge of fractions, but we plan to blend this technology-based instruction with teacher-led instruction. Therefore, a second goal of the project is to provide teachers with up-to-date information on their students' current level of understanding of fractions along with instructional tips and suggestions they can immediately apply to their pedagogy in order to help their students progress to a higher level of understanding.

In order to achieve these stated goals, we needed an accurate method to assess each student's current understanding of fractions; the HALF system will be most functional when it is able to assess what the student already knows since it cannot provide good recommendations to teachers without an accurate assessment of student progress. We discovered that while a few high-quality, well-tested fractions assessments do exist (e.g. Burns, 2008; TERC, 2008; University of Chicago School Mathematics Project, 2004), none of these assessments were a good fit for the range of curriculum covered by the HALF system, nor did they align well with the instructional model for HALF which uses visual representations of fractions to explicitly link and bridge procedural understanding with conceptual understanding.

To that end, we began to develop a diagnostic assessment of fractions knowledge. These assessment items were developed via a framework of learning objectives (Table 1) that was created using previous theoretical and empirical research about fractions (e.g., Cramer, Post, & delMas, 2002; Kieren, 1988; Moss & Case, 1999) along with mathematics content standards and benchmarks from several national organizations (e.g., Common Core State Standards Initiative, 2010; NMAP, 2008). While we currently have a developed item bank of 243 items that can be administered in one of several static assessments, our goal is to eventually create a computer-adaptive assessment of fractions that can both assess a student's current understanding of fractions as well as diagnose any specific gaps in knowledge. However, in order to accomplish this goal, we need to increase our understanding of the performance of each item and attempt to determine the reason an item is easier or harder for a student. This understanding will not only inform the progression of the learning environment system but will allow us to accurately develop items within a particular content area and item difficulty.

**Purpose / Objective / Research Question / Focus of Study:**

When examining the difficulty parameter estimates for the developed fractions items in

the item bank, we observed that there were conceptual trends related to item difficulty that crossed content areas. That is, while there was a definite progression of item difficulty from easy to hard along the stated learning objectives, some trends appeared to overlay this progression and make an item easier or harder regardless of the particular content. For example, items that used a triangle as a visual representation appeared to be consistently more difficult than items using other shapes. The purpose of this study was to determine if the observed difference in item difficulties for these trends was significant.

**Setting:**

The school sites whose students participated in the diagnostic assessment item pilot testing include nine different schools that represent a broad range of possible school settings in order to allow for more generalizable results. These school sites currently include 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 currently in the sample are located in Middle or Eastern Tennessee, although we do have plans to extend data collection to other locations.

**Population / Participants / Subjects:**

The sample includes 880 fifth- and sixth-grade students drawn from the schools described above. We are still in the process of obtaining demographic information for some members of the sample and so exact percentages of demographic categories are not available at this time, although the general characteristics of the sample are as follows. Roughly equal numbers of males and females participated. The students represent a range of math ability, as measured by standardized test results in math. 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. As in the selection of school sites, we attempted to collect data from as broad of a range of fifth- and sixth-grade students as possible; no students were systematically excluded.

**Intervention and Data Collection:**

The assessment items were administered to intact classrooms of fifth- and sixth-graders via a platform-independent computer software that delivers a set of fractions assessment items as an extensible web portal. The assessment was delivered using a classroom set of iPad tablets and a closed network server operating from a laptop computer in the classroom. Student responses were collected by the server via the closed network and saved in a database. After the class had completed the assessment, standardized database tools (e.g., SQL scripts) were used to extract and aggregate the students' responses for further analysis.

The items were developed in four iterative rounds. They were created to align with a particular content area in the learning objective framework as well as in response to the performance of items developed in previous rounds. The item bank is currently comprised of 243 developed and tested items. The items were administered in one of 14 versions of the exam, each of which was comprised of a mix of new items and previously tested anchor items. All items were multiple-choice and were scored dichotomously. Each student was given one version of the assessment, and the majority of students completed the assessment in under 30 minutes.

**Research Design and Analysis:**

The data were analyzed using the Winsteps computer program (Linacre, 2012) and fit using a dichotomous Rasch model (Rasch, 1960). This model is a probabilistic function in which a single parameter, the item difficulty *b*, is estimated for each item in the model. The dichotomous Rasch is defined as:

$$p\left(X_{ij} = 1, \theta_i, b_j\right) = \frac{e^{\theta_i - b_j}}{1+e^{\theta_i - b_j}} \quad (1)$$

where $x_{ij}$ is the response of examinee *i* on item $j$, $\theta_i$, is the ability estimate of examinee *i*, and $b_j$ is the difficulty estimate of item $j$. The results were examined via a table of *b* estimates (Table 2) and graphically depicted via a set of strip charts (Figure 1). From here, categories of items that exhibited an unusual or unexpected distribution of difficulties were examined to determine which item characteristics could be causing the differences in difficulties.

Via this visual examination, we observed a number of general trends in the data that crossed content areas. One such trend is that assessment items that use a fraction in reduced form are easier than those that use fractions in non-reduced form (e.g. 3/4 is an easier fraction to work with than 6/8). Items that present fractions that are greater than one as improper fractions are more difficult than ones that use the mixed number format. Items that present material using only symbolic notation are more difficult than those that use both symbols and visual representation. With respect to items that include visual representations of fractions, those that use a triangle shape are more difficult for students than those that use another shape (e.g. circle or square). Additionally, items that use a number line as the visual representation are generally harder than items that ask the same type of question using other visual representations or using only symbolic notation. Finally, students appear to have more difficulty identifying the whole shape when given a fractional part than they have identifying a fractional part given the whole.

The difference between each pair of groups was tested using a Welch's *t* test, which was chosen both because it does not assume equal variances between the comparison groups and because it is an appropriate test for comparing groups of item difficulty estimates (J. M. Linacre, personal communication, September 26, 2012). The items selected for analysis were chosen because they created the most directly comparable groups. For example, not every question that used a number line was included in the sample; only those number line items that had an equivalent non-number line questions were selected. An additional limit on the data is that some groups (e.g. items using triangles in the visual representation) were not well-represented in the item bank and so had a small *n*.

**Findings / Results:**

The descriptive statistics, test statistics, and 95% confidence intervals for all of the comparisons are shown in full in Table 3. Of the six *t* tests performed, four showed a significant difference between the groups. The most striking result was the comparison between items that presented a fraction that was greater than one as an improper fraction and those presented as mixed numbers ($t(27) = 7.902$, $p < .001$). This is contrary to our initial theory laid out in the learning objectives, in which improper fractions and mixed numbers were hypothesized to be of equivalent difficulty due to both being representations of fractions greater than one. However, our theory regarding this result is that while improper fractions require a relatively sophisticated understanding of the relationship between the parts of a fraction, a mixed number is easier understand because students are able to divide the parts into two components, that of a whole number and a proper fraction.

Another significant finding is the comparison between items that use both symbolic and visual representations and those that just use symbolic language ($t(30) = -3.793$, $p < .001$), in which students found the symbol-only items to be harder. This is likely because students who did not understand the symbolic language were able to use the visual representation in the question to figure out the answer. However, this finding highlights a challenge in future question development. The items that included both visual representations and symbolic were designed to test a student's ability to translate between the forms of fractions representation; however, it appears that students were instead able to bypass the translational aspect of the question and simply use the information that was easiest for them to understand.

Two of the comparisons were not significant, those in which number line questions were compared to non-number line items ($t(54) = .348$) and the comparison of part-to-whole and whole-to-part items ($t(16) = -.936$). The first non-significant result could be due to an imperfect match in items between the number line and non-number line groups. Our observations regarding the relative difficulty of number line items did not occur until after the most recent round of item development, so future research will include items that are more directly comparable to the number line questions. In the second non-significant result, items in which a student was asked to identify a whole from a fractional part were more difficult, but not significantly so, than the reverse (identifying a fractional part given the whole object). This result is likely due to the large range of possible values for each group.

**Conclusions:**

These findings are useful in multiple ways. First, they will allow us to more accurately develop items of a desired difficulty within a content domain. In our next round of item development, we will attempt to confirm these results by developing items with a specific desired item difficulty. These results are also useful in our work in improving fractions pedagogy, as they allow us to make general recommendations to teachers regarding the order in which to present material. For example, by knowing that items using triangles are more challenging for students, a teacher could introduce a new topic using circle and square models and move on to triangles when the students are ready to challenge their understanding.

These results come with an important caveat. The item difficulty estimates are derived from the performance of actual fifth- and sixth-grade students in the U.S. These empirically-derived estimates describe what the students in the sample found to be harder, not what is necessarily an inherently more difficult concept. The degree to which the item was challenging could be an artifact of poor quality prior instruction in fractions, language barriers, or even cultural pedagogical norms of how number sense is developed. However, instruction does not take place in a vacuum; regardless of how and in what order the theory states students should learn fractions, it is valuable to both good pedagogy and good research to understand how they actually do learn fractions.

# Appendices

## Appendix A. References


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.

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.

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

Gamoran, A., & Hannigan, E. C. (2000). Algebra for everyone? Benefits of college-prepratory mathematics for students with diverse abilities in early secondary school. *Educational Evaluation and Policy Analysis, 22*(3), 241-254. doi:10.2307/1164242

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

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.

Kloosterman, P., & Lester, F. (2004). *Results and interpretations of the 1990 through 2000 mathematics assessments of the National Assessment of Educational Progress*. Reston,VA: The National Council of Teachers of Mathematics.

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

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

Mullis, I., Dossey, J., Owen, E., & Phillips, G. (1991*). The state of mathematics achievement: NAEP's 1990 assessment of the nation and the trial assessment of the states* (NCES Publication No. 91-1259). Washington, DC: National Center for Education Statistics.

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.

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

United States Department of Education (1997). *Mathematics equals opportunity*. Washington, DC: Department of Education. Retrieved from ERIC database. (ED415119).

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

**Appendix B. Tables and Figures**

Table 1

*Learning Objectives for the HALF Program*

**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

*Subset of Item Difficulty (b) Estimates.*

| Item | *n* | *b* | *se* | *M* | *s* |
|---|---|---|---|---|---|
| Objective 1b.01 | 100 | -1.46 | 0.32 | 0.87 | 0.34 |
| Objective 1b.02 | 252 | -2.69 | 0.28 | 0.94 | 0.24 |
| Objective 1b.03 | 146 | -2.63 | 0.44 | 0.96 | 0.20 |
| Objective 1b.04 | 98 | -2.15 | 0.42 | 0.93 | 0.26 |
| Objective 1b.06 | 147 | -2.32 | 0.29 | 0.89 | 0.31 |
| Objective 1b.08 | 147 | -2.92 | 0.35 | 0.93 | 0.25 |
| Objective 1d.01 | 155 | 1.57 | 0.22 | 0.50 | 0.50 |
| Objective 1d.02 | 162 | -0.25 | 0.26 | 0.80 | 0.40 |
| Objective 1d.03 | 98 | 0.41 | 0.26 | 0.64 | 0.48 |
| Objective 1d.04 | 65 | 2.53 | 0.34 | 0.46 | 0.50 |
| Objective 1d.05 | 146 | -1.25 | 0.28 | 0.88 | 0.33 |
| Objective 1d.07 | 147 | 1.91 | 0.25 | 0.29 | 0.45 |
| Objective 1d.08 | 167 | -0.35 | 0.19 | 0.66 | 0.47 |
| Objective 1d.09 | 147 | -1.69 | 0.26 | 0.83 | 0.38 |
| Objective 1d.10 | 167 | -0.94 | 0.2 | 0.75 | 0.43 |
| Objective 1f.01 | 68 | -2.25 | 0.66 | 0.96 | 0.21 |
| Objective 1f.02 | 100 | -3.21 | 0.6 | 0.97 | 0.17 |
| Objective 1f.02 | 69 | -1.61 | 0.5 | 0.93 | 0.26 |
| Objective 1f.03 | 101 | -2.66 | 0.53 | 0.96 | 0.20 |
| Objective 1f.03 | 65 | -3.62 | 1.05 | 0.98 | 0.12 |
| Objective 1f.04 | 242 | -2.37 | 0.25 | 0.91 | 0.28 |
| Objective 1f.04 | 61 | -0.52 | 0.48 | 0.90 | 0.30 |
| Objective 1f.05 | 67 | -1.12 | 0.56 | 0.93 | 0.26 |
| Objective 1f.06 | 167 | -3.18 | 0.38 | 0.95 | 0.21 |
| Objective 1f.06 | 65 | 1.71 | 0.37 | 0.58 | 0.49 |
| Objective 1f.07 | 147 | -2.92 | 0.35 | 0.93 | 0.25 |
| Objective 1g.01 | 68 | 1.26 | 0.29 | 0.59 | 0.49 |
| Objective 1g.02 | 69 | 1.02 | 0.32 | 0.62 | 0.48 |
| Objective 1g.03 | 65 | 0.68 | 0.3 | 0.58 | 0.49 |
| Objective 1g.04 | 61 | 0.55 | 0.35 | 0.79 | 0.41 |
| Objective 1g.05 | 67 | 0.36 | 0.38 | 0.79 | 0.41 |
| Objective 1g.06 | 65 | -0.22 | 0.42 | 0.83 | 0.37 |

Table 2 *continued*

| Item | *n* | *b* | *se* | *M* | *s* |
|---|---|---|---|---|---|
| Objective 2c.01 | 160 | -0.36 | 0.21 | 0.74 | 0.44 |
| Objective 2c.02 | 162 | -1.57 | 0.32 | 0.92 | 0.27 |
| Objective 2c.03 | 98 | -0.14 | 0.28 | 0.72 | 0.45 |
| Objective 2c.04 | 146 | 0.73 | 0.26 | 0.63 | 0.48 |
| Objective 2c.06 | 147 | 0.44 | 0.2 | 0.51 | 0.50 |
| Objective 2c.07 | 167 | -1.15 | 0.22 | 0.78 | 0.42 |
| Objective 2c.08 | 147 | 0.15 | 0.21 | 0.56 | 0.50 |
| Objective 2c.09 | 232 | 0.21 | 0.22 | 0.62 | 0.49 |
| Objective 4ca.01 | 242 | -1.27 | 0.18 | 0.81 | 0.40 |
| Objective 4ca.02 | 101 | -1.69 | 0.38 | 0.91 | 0.28 |
| Objective 4ca.03 | 98 | -0.94 | 0.3 | 0.83 | 0.38 |
| Objective 4ca.04 | 147 | -0.73 | 0.21 | 0.70 | 0.46 |
| Objective 4ca.05 | 146 | -0.82 | 0.26 | 0.84 | 0.37 |
| Objective 4ca.13 | 167 | -0.66 | 0.2 | 0.71 | 0.46 |
| Objective 4ca.14 | 167 | -1.15 | 0.21 | 0.78 | 0.42 |
| Objective 4cb.06 | 303 | 1.84 | 0.16 | 0.37 | 0.48 |
| Objective 4cb.07 | 100 | 1.36 | 0.28 | 0.47 | 0.50 |
| Objective 4cb.09 | 98 | 1.93 | 0.29 | 0.41 | 0.49 |
| Objective 4cb.10 | 68 | 3.65 | 0.41 | 0.22 | 0.41 |
| Objective 4cb.12 | 147 | 1.04 | 0.21 | 0.41 | 0.49 |
| Objective 4cb.13 | 65 | 2.58 | 0.36 | 0.25 | 0.43 |
| Objective 4cb.14 | 61 | 3.71 | 0.35 | 0.28 | 0.45 |
| Objective 4cb.15 | 67 | 3.19 | 0.31 | 0.34 | 0.47 |
| Objective 4cb.16 | 147 | 2.13 | 0.24 | 0.27 | 0.44 |
| Objective 4cb.17 | 167 | 2 | 0.22 | 0.29 | 0.45 |
| Objective 4cb.18 | 65 | 3.51 | 0.35 | 0.32 | 0.47 |
| Objective 5a.01 | 68 | 0.82 | 0.3 | 0.66 | 0.47 |
| Objective 5a.02 | 69 | 0.38 | 0.32 | 0.72 | 0.45 |
| Objective 5a.03 | 65 | -0.19 | 0.31 | 0.74 | 0.44 |
| Objective 5a.04 | 61 | 1.11 | 0.32 | 0.70 | 0.46 |
| Objective 5a.05 | 67 | 0.91 | 0.32 | 0.72 | 0.45 |
| Objective 5a.06 | 65 | 0.47 | 0.4 | 0.75 | 0.43 |

Table 3

*Descriptive Statistics, Welch t test Statistics, and 95% Confidence Intervals for Comparisons*

| Group | *n* | *M* | *s* | *t* | *se* | 95% CI (lower) | 95% CI (upper) |
|---|---|---|---|---|---|---|---|
| Used triangle (1) | 4 | 1.135 | 1.896 | 3.997** | 0.359 | 0.597 | 2.270 |
| Did not use triangle (0) | 29 | -0.297 | 1.406 | | | | |
| Used number line (1) | 26 | 0.197 | 1.001 | 0.348 | 0.257 | -0.426 | 0.605 |
| Did not use number line (0) | 37 | 0.107 | 1.010 | | | | |
| Given whole, what is the part (1) | 9 | -0.406 | 1.322 | -0.936 | 0.664 | -2.030 | 0.788 |
| Given part, what is the whole (0) | 9 | 0.216 | 1.490 | | | | |
| Item required reducing fractions (1) | 7 | 0.483 | 0.708 | 4.352* | 0.597 | 1.017 | 4.179 |
| Item did not require reducing fractions (0) | 4 | -2.115 | 1.067 | | | | |
| Item used improper fractions (1) | 17 | 2.159 | 1.134 | 7.902*** | 0.350 | 2.050 | 3.489 |
| Item did not use improper fractions (0) | 12 | -0.610 | 0.752 | | | | |
| Item used both model and symbol (1) | 12 | -0.615 | 0.684 | -3.793*** | 0.298 | -1.741 | -0.522 |
| Item used symbolic notation only (0) | 21 | 0.517 | 1.026 | | | | |

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

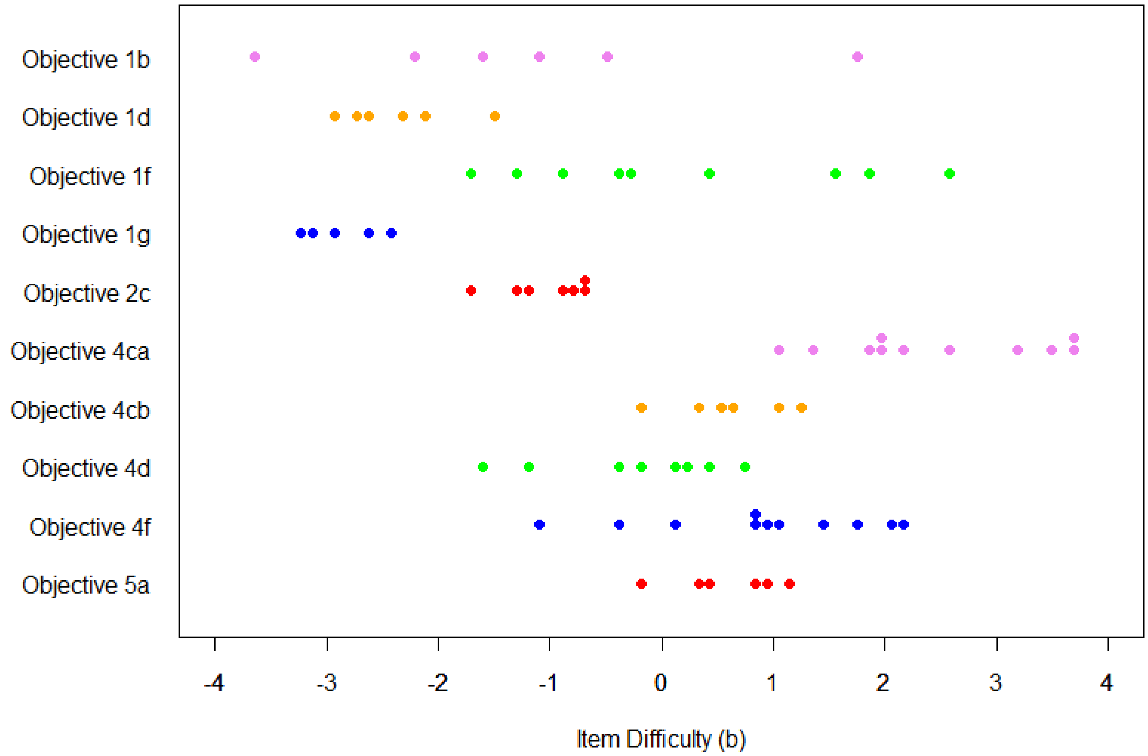


*Figure 1*. Strip chart showing a subsample of item difficulty estimates. According to the learning objectives (Table 1), the *b* estimates of the easiest (Objective 1) questions should have been clustered on the left and the *b* estimates of the hardest (Objective 5) questions clustered on the right.