# MEASURING CHILDREN'S PROPORTIONAL REASONING, THE "TENDENCY" FOR AN ADDITIVE STRATEGY AND THE EFFECT OF MODELS

Christina Misailadou and Julian Williams

University of Manchester

*We report a study of 10 -14 year old children's use of additive strategies while solving ratio and proportion tasks. Rasch methodology was used to develop a diagnostic instrument that reveals children's misconceptions. Two versions of this instrument, one with "models" thought to facilitate proportional reasoning and one without were given to a sample of 303 children. We propose a methodology for examining systematically the pupils' additive errors, their effect on ratio reasoning and how contingent on "model" presentation this is. First we provide a measure on which pupils, item-difficulty and additive errors can be located. We then construct a new measure, which we name "tendency for additive strategy". Finally, we find that the presence of "models" affects this new measure and draw inferences for choices of items in assessment and teaching.*

## INTRODUCTION

This study builds on previous work on children's misconceptions while solving ratio and proportion tasks and especially on their use of an "additive strategy" to obtain an answer.

The additive strategy is the most commonly reported erroneous strategy in the research literature related to ratio and proportion. When using this strategy to solve a ratio item, "the relationship within the ratios is computed by subtracting one term from another, and then the difference is applied to the second ratio." (Tourniaire & Pulos 1985, p.186)

In this study we aim to contribute to teaching by developing an instrument that can help teachers diagnose their pupils' misconceptions, including the use of the additive strategy. Twenty-four, missing value items were used to construct the instrument. Some of these items have been adopted with slight modifications of those used in previous research and others have been created based on findings of that research. (CSMS 1985, Lamon 1993, Tourniaire 1986, Cramer, Bezouk & Behr 1989, Kaput & West 1994, Singh 1998) All of the items were selected having as a criterion their diagnostic potential as reported in the above studies. This is their potential to provoke a variety of responses from the pupils, including errors stemming from well known misconceptions. Furthermore, we tried to use a variety of problems as far as "numerical structure" and "context" is concerned. As a result of this selection, errors indicative of common and frequent misconceptions such as the additive strategy were expected to occur. In addition, we hoped that less frequent misconceptions or even ones that are not mentioned in the research literature would also occur.

Finally, two versions of the instrument were tested (both of these versions can be seen in full on the web at http://www.education.man.ac.uk/lta/cm/). The first version ("W") contains all the twenty-four items presented as mere written statements. The second version ("P") contains the same items but this time most of them are supplemented by "models" thought to be of service to children's proportional reasoning. These models involve pictures, tables and double number lines. Our purpose was to compare the difficulty of the parallel items for the children and to spot differences in the strategies used in each type of items.

This paper reports the results from the scaling of both versions of the instrument and focuses particularly on the occurrence of the additive strategy. The scaling of the instrument provides a measure on which pupils, item difficulty and additive errors can be located. We present the results for its non-model form. Based on these results, we construct a new measure which we name "tendency for additive strategy" (again for the non model form). Finally we examine the "model effect" on the tendency for additive strategy.

## METHOD

In order to be able to administer more items to the same sample of pupils, each version of the test consisted of two separate test forms with common linking items. Thus, Test W was divided in Test W1 and Test W2. Test W1, designed to be easier, consisted of sixteen items and Test W2 had the same number of items, but was designed to be more difficult. Eight of the items were common for both tests. Exactly the same pattern applies for tests P1 and P2 into which Test P was divided. Finally, we equated Test W1 and P1 through common items and we did the same for Test W2 and Test P2 in order to be able to compare the difficulty of the parallel items for the children.

The tests were given to a sample of 303 pupils aged 10 to 14 years old from 4 schools in the north west of England. Before administering the tests to the pupils, their teachers were asked to comment on the suitability of the items for their classes. They found that the items were generally acceptable for the pupils' age and viewed them as valid assessment of the curriculum they are teaching. They commented though, on the difficulty of the items 3Paint and 6OnionSoup.

For each item of the test, all the pupils' answers, correct and erroneous, were coded and the results were subjected to a Rasch analysis in the usual way using the program Quest. (For a summary of this method see Williams and Ryan, 2000: the Rasch scaling is the modern stochastic development of the Guttman scaling model used in the CSMS studies reported by Hart, 1981). This analysis allowed us to scale the most common errors for each item with its difficulty in the W and P form. The result was a single difficulty estimate for each item and an ability estimate for each child consistent with the Rasch measurement assumptions. Item "3Paint" fell outside a model infit statistic value of 1.3 (see Wright & Stone, 1979) reflecting the difficulty of this item for the sample.

Finally, we were able to validate the interpretation of significant misconceptions because for each item in the test, there were specific instructions to the children not only to write an answer but an explanation for it as well. Furthermore, in addition to the test analyses, we drew on structured clinical interviews with 13 children and structured small group interviews with 63 children about the test items. These interviews allowed us to validate the items and to confirm our interpretations about the strategies that were used to solve the items.

## RESULTS

**Scale of performance and additive errors**

Several factors that are supposed to make ratio problems more susceptible to additive errors are mentioned in previous research studies. Some reports draw attention to numerical factors (e.g. Bell, Costello and Küchemann, 1985, mention as such a factor, the appearance, in a problem, of fractions other than halves) and others point out the context factors (e.g. Kaput and West, 1994, mention the "geometric figures" problems as the most vulnerable to additive errors). Since we tried to have in our instrument problems that are as varied as possible in numerical structure and context we decided to have a more systematic approach to the kind of additive errors that pupils make and to the kind of problems, in terms of numerical and context characteristics, that provoke an additive approach.

Table 1 below, consists of four columns. The numbers on the first column, beginning from the left, represent the logit scale on which items and cases are calibrated. The second column (the "Xs") presents the distribution of the cases (pupils) according to their ability estimate. Here by the term "ability" we mean the performance of the pupil in this particular test and not a general mathematical ability. The third one (where each name represents a different question of our instrument) shows the distribution of the items (on the same scale) in reference to their difficulty estimates. (The names and the corresponding items can also be seen at the above web address) Finally, the last column presents the mean abilities of the pupils that produced particular additive errors. We name each of these errors by using the prefix "ad" and the name of the corresponding item. For example, by "Ad4OnionSoup" we indicate the mean ability of the pupils that gave an erroneous answer to the item "4OnionSoup", as a result of the additive strategy. The additive errors listed in the table are most likely to be made by children at the ability adjacent or below. We have included only errors that occurred on more than 3% of the scripts on the grounds that one might expect to see one such occurrence in a random classroom of 30 children.

From the table below, it can be inferred that the mean ability of many of the pupils that gave additive errors is quite close to the mean ability of our sample. Afantiti-Lamprianou &Williams (2002) have demonstrated the usefulness of a different technique for scaling errors with case abilities, for a "representativeness tendency" measure as a diagnostic measure for probability items. Accordingly, we considered building a "tendency for additive strategy" measure as a diagnostic measure of tendency to inappropriately apply this strategy.

```
      Pupils' Ability      Items' Difficulty                    Additive errors

  |            XX  | 3Paint
  |              |
  |              |
4.0|              |
  |              |
  |              |
  |            X  |
  |              |
  |              |
3.0|              |
  |            XX  |
  |              |
  |           XXXX  |
  |              | 6OnionSoup, 3Campers
2.0|              | 1Rectangles
  |            X  | PrintingPress
  |              | 1Paint
  |        XXXXXXXXXX  |
  |         XXXXXXXX  |
  |          XXXXX  | MrShort, Biscuits
1.0|            XXX  | OfficeBill
  |          XXXXXX  | 2Campers
  |         XXXXXXXX  | 1Campers
  |        XXXXXXXXX  | 5OnionSoup
  |       XXXXXXXXXXX  |
 .0|           XXXXX  | 2Paint
  |              |
  |            X  |
  | XXXXXXXXXXXXXXXXXXXXXXXXXX  |                        Ad3Paint
  |              |                                       Ad1Paint, AdMrShort
  |    XXXXXXXXXXXXXXXXXX  | Books'Price                Ad1Rectangles
-1.0| XXXXXXXXXXXXXXXXXXXXXXXXXX  | 3OnionSoup            Ad3Campers
  |          XXXXXX  | Party, 2Rectangles               Ad3Oni, Ad5On, Ad1Cam, Ad2Cam,AdPri
  |    XXXXXXXXXXXXXXXXX  | 4OnionS, Fruits'Pric        Ad2OnionSoup, Ad2Paint,AdBookPrice
  |         XXXXXXX  |                                  Ad1Eel, Ad2Eel, AdParty, Ad6OnionS
  |              |
-2.0|            XXXXXX  | 2OnionSoup, 2Eels             Add4OnionSoup
  |     XXXXXXXXXXXXXXX  |
  |         XXXXXXX  | 1Eels
  |              |
  |          XXXXX  | Class
  |     XXXXXXXXXXXXXXX  |
-3.0|           XXXXXXX  |
  |            XX  | Reading
  |              | 1OnionSoup
  |              |
  |           XXXX  |
-4.0|              |
  |              |
  |              |
  |              |
  |            XX  |
  |              |
-5.0|              |
```

Table 1: Scale of performance and additive errors

## The "tendency for additive strategy"

A second Rasch analysis was run using Quest. One mark was given only for the answer that indicated the use of an additive strategy and no marks were given for any other responses. The result was a single scale of items that can be seen in Table 2 below (none of the items fell outside a model infit value of 1.3).

Children that are higher up in this scale are more likely to use the additive strategy and items that are higher up in the same scale are the least likely to provoke additive errors-in these items only the pupils with a strong "tendency for additive strategy" made such errors.

```
            XX  |
                | OfficeBill
                |
                |
 2.0             | Reading
                |
                | 2Rectangles
                | 3OnionSoup
                |
                |
             X  | 2OnionSoup, 4OnionSoup
             X  | 6OnionSoup
 1.0             |
             X  |
                | Books'Price, Party
                | 1Eels, 2Eels
            XX  |
                |
                |
            XX  |
  .0        XXXXXXX  | PrintingPress
    XXXXXXXXXXXXXXXXXX  |
                |
          XXXXXX  |
                |
                | 2Paint
      XXXXXXXXXXXXXX  | 1Paint
-1.0             |
                |
         XXXXXXXXX  | 5OnionSoup, 1Rectangles            | Items
    XXXXXXXXXXXXXXXXXX  |                         | most
                | 2Campers                   | likely
                | 3Paint                    | to
         XXXXXXXX  | 1Campers                  | provoke
   XXXXXXXXXXXXXXXXXXXXX  |                          | the use
-2.0      XXXXXXXXXXXXX  | 3Campers                   | of the
                |                         | additive
           XXXX  | MrShort                   | strategy
                |
                |
                |
       XXXXXXXXXXXX  |
        XXXXXXXXXXX  |
-3.0             |
                |
                |
                |
             X  |
      XXXXXXXXXXXXXXX  |
                |
```

Table 2: Item and Person Estimates of the "tendency for additive strategy"

The outcome of this analysis is a measure of "tendency for additive strategy" for each person, which naturally correlates negatively with their "ratio reasoning ability" as measured previously (rho= -0.2). The regression equation was found to be to 2 SF:

Ratio reasoning ability = -0.28 x (Tendency for additive strategy) - 1.5

According to the regression analysis of our data 4 % of the variability in "ratio reasoning ability" is accounted for by the "tendency for additive strategy". It is interesting to note that this new measure diagnosis a pupils' weakness but further research is needed for the rest of the variance to be explained. From inspection of the frequency of other errors though, we know that they will not account for much more variance than the "additive strategy" which is the strongest.

By taking into account the item analysis results for observed responses provided by Quest we distinguished on Table 2 the seven items of our test that most frequently provoke additive responses.

By inspection, these items indicate that the decisive factor for the difficulty of a ratio problem is its numerical structure. Five of these items are some of the most difficult numerically items of the instrument: neither the scalar nor the functional relationships in the given proportions are integer ratios and the answers cannot be obtained by a simple multiplication or division by an integer. Only the item "2Campers" has an easy numerical structure and consequently we assume that it is the "sharing context" that makes this item prone to additive errors.

**Effect of models on "ratio reasoning ability" and "tendency for additive strategy"**

As we have mentioned, we have created two versions of our instrument: one with and one without models. The table below (Table 3) presents the Rasch analysis estimates for the items of our instrument that were presented in both of these versions.

| | Ratio reasoning ability estimates | | | Tendency for additive strategy estimates | | |
|---|---|---|---|---|---|---|
| Item's name | Difficulty of the item without models | Difficulty of the item with models | Difference in logits | "Additive estimate" of the item without models | "Additive estimate" of the item with models | Differ. in logits |
| 1Paint | 1.70 | -.69 | -2.39 | -.76 | -1.08 | -.32 |
| Reading | -3.22 | -3.17 | .05 | 2.03 | 1.55 | -.48 |
| 2Paint | .11 | -.84 | -.95 | -.70 | -.86 | -.16 |
| Books' Price | -.81 | -.37 | .45 | .70 | .52 | -.18 |
| 1Campers | .65 | 1.68 | 1.03 | -1.73 | -1.38 | .35 |
| 2Campers | .76 | 1.25 | .49 | -1.40 | -1.38 | .02 |
| Class | -2.72 | -1.76 | .96 | No score | No score | N.A. |
| Party | -1.23 | -.88 | .35 | .70 | .63 | -.07 |
| Fruit's Price | -1.32 | -.88 | .44 | No score | 3.05 | N.A. |
| 3Paint | 4.69 | 3.45 | -1.24 | -1.52 | -1.51 | .01 |
| 3Campers | 2.19 | 2.29 | .10 | -1.91 | -1.57 | .34 |
| Printing Press | 1.85 | 1.09 | -.76 | -.03 | .99 | 1.02 |
| Biscuits | 1.19 | 1.09 | -.10 | No score | No score | N.A. |

Table 3: Comparison of the parallel items

All the estimates are given in logits. In the first scale, which measures "ratio reasoning ability" the difficulty estimates start from negative to positive logits for easier to more difficult items respectively. In the second scale, which measures "tendency for additive strategy" the additive estimates start again from negative to positive logits, for more to less frequently provoking additive errors items respectively. A positive difference in column 3 indicates that the "non-model" item is harder (as designed) whereas a positive difference in column 6 indicates that the "non-model" item is less likely to encourage an additive strategy (also as designed).

We note that in nearly half of these items the "non-model" item is actually easier than the model version, and these include the easiest items and the items for which the model does not help avoidance of additive strategy. From the seven most interesting "additively" items, four (the ones highlighted in Table 3) were presented in a model as well as a non-model version. For all of them the model was a pictorial representation of their data.

We have already reported (Misailidou & Williams, 2002) that the addition of pictures in each of those items affected the kind and the frequency of the strategies that pupils had employed. As we can see on the table, the pictorial version of the item "3Paint" was easier than the other one whereas for all three of the "Campers" items the pictorial version was more difficult than the other one!

On the other hand, by looking at the "tendency for additive strategy" for all the four items we realize that for each item the pictorial version is located higher up the scale than the "without the model" version. It seems that, with the supplement of pictures, these additive errors became less frequent. We believe that, although the addition of a pictorial representation to a ratio item does not always make it "easier" for the children, it could decrease the item's potential to trigger additive errors.

Generally, it seems that in our attempt to design "models" which support ratio reasoning, and the avoidance of additive strategies in particular, we succeeded best in the pictorial designs. We guess that the other "models" need to be taught to children and not just presented to them. Thus we see the significance of the new measure as a means of understanding the "model" effect: use of the difficulty of the model and non-model items alone is not so helpful.

## CONCLUSIONS AND DISCUSSION

Our aim was to complement what has already been reported on the children's inappropriate use of additive strategies in responding to test questions which are relevant to their curriculum. We have developed two scales which measure "ratio reasoning ability" and "tendency to additive strategy" and both scales contain "model" and "non-model" parallel items. We found that the additive tendency accounts for only a small proportion of children's problems with ratio.

The influence of presentation on item difficulty and on additive tendency is strong and Table 3 suggests that items might be chosen selectively to provoke or avoid conflict between children's different responses.

While most of the particular items we have used for our instrument are not new, the development, validation and calibration of the measures around the additive strategy is. We have demonstrated how these tools can be used for research purposes but we also believe in their importance on teaching practice and teacher education as well. We suggest that the knowledge that teachers would collect from these scales might enrich their pedagogical content knowledge about ratio and proportion (in a manner discussed in Williams & Ryan, 2001) and thus help them improve their classroom practice.

We gratefully acknowledge the financial support of the Economic and Social Research Council (ESRC), Award Number R42200034284

## References

Afantiti-Lamprianou, T. & Williams, J. (2002). A developmental scale for assessing probabilistic thinking and the tendency to use a representativeness heuristic. In A.Cockburn & E. Nardi (eds.) *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education (PME)*. Norwich: University of East Anglia.

Bell, A. Costello, J. & Küchemann, D. (1985). *Research on Learning and Teaching*. Windsor: NFER-NELSON.

Cramer, K. Bezuk, N. & Behr, M. (1989). Proportional Relationships and Unit Rates. *Mathematics Teacher,* 82(7), 537-544.

Hart, K. (1981). *Children's Understanding of Mathematics, 11-16*. London: John Murray

Hart, K. Brown, M. Kerslake, D. Küchemann, D. & Ruddock, G. (CSMS). (1985). *Chelsea Diagnostic Mathematics Tests. Teacher's Guide*. Windsor: NFER-NELSON.

Kaput, J. & Maxwell-West, M. (1994). Missing-Value Proportional Reasoning Problems: Factors Affecting Informal Reasoning Patterns. In G. Harel, & J.Confrey, (eds.) *The Development of Multiplicative Reasoning in the Learning of Mathematics*. Albany: State University of New York Press.

Lamon, S. (1993). Ratio and Proportion: Connecting Content and Children's Thinking. *Journal for Research in Mathematics Education,* 24(1), 41-61.

Misailidou, C. & Williams, J. (2002). Diagnostic Teaching and Ratio: The Effect of Pictorial representations on Strategies. In A. Cockburn & E. Nardi (eds.) *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education (PME)*. Norwich: University of East Anglia.

Singh, P. (1998). *Understanding the Concepts of Proportion and Ratio among Students in Malaysia*. PhD Thesis. The Florida State University.

Tourniaire, F. & Pulos, S. (1985). Proportional Reasoning: A review of the literature. *Educational Studies in Mathematics*, 16(2), 181-204.

Tourniaire, F. (1986). Proportions in Elementary School. *Educational Studies in Mathematics,* 17(4), 401-412.

Williams, J. & Ryan, J. (2000). National Testing and the Improvement of Classroom Teaching: Can they Coexist? *British Educational Research Journal,* 26(1), 49-73.

Williams, J. & Ryan, J. (2001). Charting Argumentation Space in Conceptual Locales. In M. van den Heuvel-Panhuizen (ed.) Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education (PME)

Wright, B. & Stone, M. (1979). Best Test Design. Chicago: MESA Press.