Journal of Research and Reflections in Education
December 2016, Vol. 10, No.2, pp 123- 131
http://www.ue.edu.pk/jrre

Students Mathematical Problem-Solving Proficiency in Relation to Gender at Grade VI

[1] Ghazala Noureen, [2] Itrat Sheikh

[1] Assistant Professor LCWU
[2] Principal Fazaia College of Education Lahore

Mathematics is abstract but logical in nature. Many concepts in mathematics cannot be explained easily in terms of physical representations or problems related to everyday life. It is a commonly accepted opinion that mathematics as a subject favors male students. A large sample of 1500 grade VI students (public sector schools) was drawn from four districts of the Punjab province. The study sought to find out the gender differential mathematic performance of the students. MCQs based achievement test was constructed and tried out on 200 students. Items for final test were selected by keeping in view the criteria of classical test theory and item response theory. It was found that in solving problems of geometry and factorization female students perform significantly better than male. Boys and girls have shown equivalent problem solving proficiency in algebra, area and perimeter, whole number and volume and surface area. Curricula devised for these subject areas might be inappropriate at these ages. Therefore, the curriculum should be thoroughly reviewed to ensure that concepts are interlinked and concepts are developed step by step.

Key words: *problem solving, item response theory, proficiency, gender*

Introduction

As per Galileo's "the whole universe is written in mathematics language and we cannot understand it without learning its letters which are triangles, circles and other geometrical figures" (Newman, 1956, p. 733). Without these, one is wandering about in a dark warren. It is a conceptual system which fulfills the inner requirements of a human. It is the general study of numerals, space, shapes, quantity and the measurements. Mathematics is also called the queen of sciences. It is a conceptual system seeking to describe the world around by considering number, space, shape, quantity and measurement. It not only is valuable in describing the physical world but is also used increasingly in the biological, medical and social world, extending its influence into descriptions of societies, behaviour and global patterns. Because of its great importance in so many areas of life, it has enjoyed a position as a core subject since antiquity, going right back to the days of the Greek civilization over 2500 years ago.

In almost all countries, mathematics holds a key position in the curriculum at all levels. At the upper levels of secondary school education, studies in mathematics can be optional. In many countries, the uptake rates are disappointing (e.g. Al-Enezi, 2008) while in others, mathematics is highly popular (Ali, 2008). In Pakistan, mathematics is often regarded as very demanding and valuable discipline.

In spite of above mentioned importance mathematics is often described as being abstract and unrelated to life. The lack of opportunity for creativity and enjoyment is a common perception and, sometimes, students find mathematics boring. Research has shown that mathematical problems set in words (often called arithmetical or algebraic story problems) are incredibly difficult which is caused by working memory limitations and lack of problem solving skills. The learner simply does not have enough knowledge and skills to handle language ideas, mathematical ideas in problem solving.

There are several ways to assess student’s problem solving abilities. Educators and psychologists have developed different tools to assess students’ problem solving skills. The achievement of the students can be assessed by classical test theories (CTT) in which all easy and difficult items have same weightage for all students. CTT does not take into account the difficulty levels of the item when estimating learning achievement of the students. During recent decades assessment based on item response theory (IRT) models have been developed to assess student's abilities. The focus of the IRT is on the pattern responses that the respondent makes to the set of items, and it does not assume that all items on the test are parallel. The essence of the IRT is the location of the person’s ability and item difficulty on the same continuum of measure.

IRT provide guidance to develop the items and to construct the proficiency scales. IRT gives us a way to understand and interpret scores and also link student’s achievement to their latent abilities. This would be helpful if the objective of an assessment is to aware teachers about students learning outcomes (Azeem, 2009) .

IRT presents students abilities and item characteristics in single mathematical scale. Once student’s ability is shown on the scale, we can easily make inferences about the type of task students become able to perform (Reckase, 1997).

Different intellectual processes such as reading, comprehension, computational skills and reasoning can be entrenched in a single task. Multi-component IRT models can help us to understand the sub-steps within a task, bringing diagnostic closer to cognitive processes (Lai, 1998).

Mathematics and Gender

There have been many studies conducted to determine the effect of several demographic variables on student achievement. This study looked specifically the effect of gender on student’s problem solving proficiency at grade VI. The results from prior studies about the effect of gender on academic achievement are mixed.

Maccoby and Jacklin (1978) after a meta-analysis based on 1,500 cross-cultural studies stated that clear and consistent gender differences emerge after age 11, with girls becoming superior in verbal abilities and boys in mathematical and visual-spatial abilities. In 1989 Friedman conducted meta-analysis focused on studies published between 1974 and 1987. Initially no gender difference was found in mathematics achievement but with the passage of time performance of male students become greater. Later on studies supported the stereotypical perception about female inferiority in mathematics. Hyde (1996) determined from meta-analysis that the magnitude of the gender difference in mathematics ability decreased by half in two decades between 1974 and 1996 (Bezzina, 2010).

Hargreaves et al. (2008) tried to investigate the gender differential performance in “gifted and talented” 9 and 13 year olds in mathematics assessment in England. The result of this study showed that there was no significant gender difference in performance for the 9 or 13 year olds. “…attitudinal differences were found, including a seemingly commonly held stereotypical view of mathematics as boys' subject. Further findings reveal that “gifted” girls perform as well as “gifted” boys, but their confidence in the subject is lower than their performance might suggest.” According to them, their results are important since the uptake of higher-level mathematically based courses by girls in England is poor.

As we progress through the twenty-first century, studies conducted in Australia, England, Hong Kong, New Zealand, Pakistan and many other countries around the world generally show that girls are performing better than boys in mathematics at the compulsory school level (Downing et al., 2008; Foster et al., 2001; Gorard et al., 2001; Rowe and Rowe, 2002; Saeed and Bushra, 2005) and fears about underachieving boys at the compulsory school level continue to shape educational discourse (QCA, 2003). Many studies in the UK reported that the boys ‘‘laddish’’ behaviour was acting as an impediment to the progress of boys. All the research results shown above come from different countries with different sample sizes but do not offer any generalised conclusion. The achievement in mathematics with respect to gender involves many factors.

Any differences in performance and attitudes between boys and girls might be attributes to a variety of factors, including:

(a) Equal opportunity of learning for boys and girls

(b) Cultural dimensions of the society

(c) Differences in the way the brains of boys and girls handle mathematical ideas.

(d) Level of gender neutrality of the curriculum, textbooks ans assessment items.(e)

Pakistani society is male dominant and girls usually do not get very strong support from their parents and male education is considered more imperative while considering their future bread winning role. Things are changing slowly and steadily and girls are coming forward to science professions but still the ratio is not satisfactorily high. Parents' support, encouragement from schools at early years, social acceptance, equal career opportunities, mathematics curriculum designed for keeping everybody in mind regardless of social, age and gender differences, all can help girls to be more confident and comfortable with mathematics..

Research Methodology

The study was descriptive in nature aimed to investigate gender differences in relation to mathematics problem solving skills at grade six. All public and private schools of district Lahore, Faisalabad, Sargodha and Jhung were the population of the study. Seventeen schools (1500 students) of above mentioned districts were selected randomly as the sample of the study.

Table 1

Distribution of sample Schools by District

District	Number of schools	Male	Female
Lahore	5	3	2
Faisalabad	5	3	2
Sargodha	3	1	2
Jhang	3	2	1

Instrumentation

Item Response Theory is a psychometric Theory and family of associated mathematical models that relate latent trait(s) of interest to the probability of Responses to Items on the assessment. The main purpose of IRT is to create a scale for the interpretation of assessments with useful properties Scaling refers to the process by which we choose a set of rules for measuring a phenomenon, The Item Characteristic Curve (ICC) is the primary concept in IRT. An ICC is a mathematical expression that connects or links a subject's probability of success on an item to the trait measured by the set of test items. The ICC is a nonlinear (logistic) regression line, with item performance regressed on examinee ability.

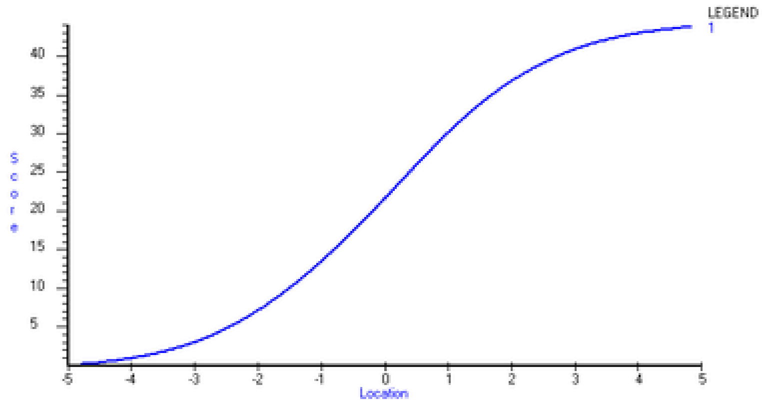


Figure 1: Test characteristic curve showing the relationship between total score on a test and person location estimate

IRT has different models depending upon item parameters as invariant over samples of examinees from the population for whom the test is intended. RASH model is used in this study, In the Rasch model, the probability of a specified response (e.g. right/wrong answer) is modeled as a function of person and item parameters. Specifically, the probability of a correct response is modeled as a logistic function of the difference between the person and item parameter.

Data analyzed using the model is usually responses to conventional items on tests. However, the model is a general one, and can be applied wherever discrete data are obtained with the intention of measuring a quantitative attribute or trait.

A proficiency test based on IRT was developed for this study. The main objective of this test was to evaluate students' problem solving proficiency in Mathematics at level VI. Firstly a proficiency test consisting 50 items was developed for piloting. This test was based on multiple choice (MCQ) questions. There are several advantages to multiple choice tests, if item writers are well trained and items are quality assured, it can be a very effective assessment technique. Multiple choice questions lend themselves to the development of objective assessment items. Multiple choice tests often require less time to administer for a given amount of material than would tests requiring written responses. This results in a more comprehensive evaluation of the candidate's extent of knowledge.

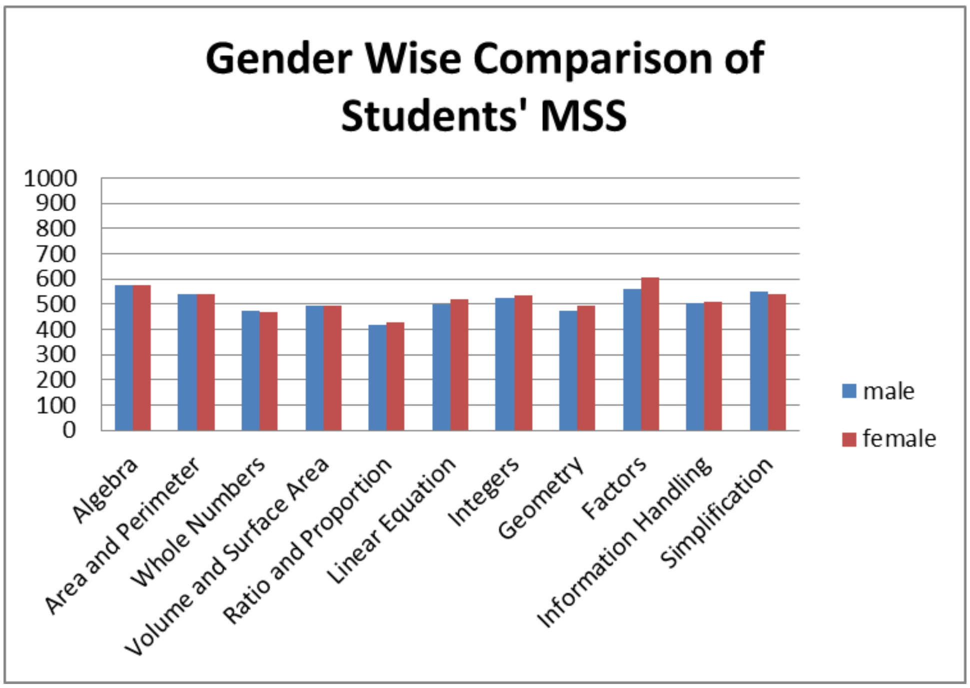
Test was connected with benchmarks and standards of Mathematics for 6^{th} class and its framework was aligned with National Assessment Examination Program (NAEP). This framework provides guideline about assessing the Mathematical problem-solving proficiency of the students. The primary data gathered by piloting was analyzed through software's ITEMEN and CONQUEST for both CTT and IRT. After pilot test the items were selected and finalized for the final data collection. Final test was comprised of 30 MCQs.

Results and Discussion

Table 2
Descriptive Statistic of Test

Strand	N	Mean	SD
Whole Numbers	1500	471.29	106.06
Integers	1500	526.55	112.91
Factors and Multiples	1500	573.67	123.82
Ratio and Proportion	1500	419.47	107.08
Algebra	1500	576.07	143.34
Area and Perimeter	1500	538.36	129.38
Geometry	1500	479.60	134.70
Volume and Surface Area	1500	493.56	109.69
Information Handling	1500	504.98	109.77
Linear Equation	1500	503.93	142.08

Mean is the mean scaled score of each student's problem-solving proficiency calibrated with an arbitrary scale 0-1000 with mean 500.



Note::Mean Square Score (MSS)

The above graph shows the performance of student in different content areas of mathematics in relation to gender

Table 3

Gender Comparisons

Strand	Gender	N	Mean	St Dev	t	P
Whole Numbers	*Male*	1069	472	106.3	0.38	n.s.
	Female	431	470	105.7		
Integers	*Male*	1069	524	109.8	1.53	n.s.
	Female	431	534	120.1		
Factors and Multiples	*Male*	1069	560	125.0	6.88	< 0.001
	Female	431	608	114.1		
Ratio and Proportion	*Male*	1069	417	112.5	1.55	n.s.
	Female	431	426	92.0		
Algebra	*Male*	1069	576	143.3	0.01	n.s.
	Female	431	576	143.7		
Area and Perimeter	*Male*	1069	538	131.1	0.20	n.s.
	Female	431	539	125.1		
Geometry	*Male*	1069	475	141.2	2.24	< 0.05
	Female	431	492	116.4		
Volume and Surface Area	*Male*	1069	493	109.7	0.25	n.s.
	Female	431	495	109.8		
Information Handling	*Male*	1069	503	109.9	1.11	n.s.
	Female	431	510	109.5		
Linear Equation	*Male*	1069	498	144.5	2.76	< 0.01
	Female	431	520	134.7		

In most areas of the curriculum, there are no gender differences in student performance, as measured by the test. However, in three areas the girls out-perform the boys: Factors and multiples, Linear equations, Geometry. It has to be noted that, with such large samples, fairly small difference in performance are found to be statistically different.

Discussion

In the area of performance in mathematics at various ages and stages, the literature seems to offer a somewhat confused picture in relation to gender. Some have found that males out-perform females (Benbow and Stanley 1983; Fennema *et al.*, 1998; Halpern *et al.*, 2007;; Mills *et al.*, 1993) while others have found the reverse (Hyde, *et al.*, 1990) while many studies found few differences. The differences are almost certainly easy to explain in terms of culture, different curricula, different ways of teaching and differences in age. In all the studies, any differences are found to be quite small. There is no intrinsic logical reason why males and females should perform differently in mathematics.

It is possible to set up curricula with male or female bias, to examine in ways that preferentially benefit either males or females. The outcomes from the study described in this paper confirm the general patterns that have been found elsewhere.

In three areas, there is a gender difference but in only one area (factors and multiple) the difference is significant at $p<0.001$.

In none of the three areas is the difference very large in terms of scores, the significance being detected simply because of large samples. This suggests that, in the context of Pakistan education at grade VI males and females perform very similarly.

In that mathematics underpins so many career options, it is vital that Pakistan enables every student to perform to their best in mathematics and, to that end, the way the curricula are planned, implemented and assessed may be important factors.

In that the females perform as well as males (and occasionally better), it is important that the females are given every opportunity to develop their skills in mathematics and to make a contribution throughout life.

References

Alexander, P. A., White, C. S., & Daugherty, M. (1997) Analogical reasoning and early mathematics learning In L. D. English (Ed.) Mathematical reasoning, Analogies, metaphors, and images (pp. 117–147). Mahwah, NJ: Erlbaum.

Alenezi, D.F. (2008) *A Study of Learning Mathematics Related to some Cognitive Factors and to Attitudes*, PhD Thesis, Glasgow: University of Glasgow. [http://theses.gla.ac.uk/333/]

Ali, A.A and Reid,N.(2012) Understanding Mathematics: Some Key Factors, *European Journal of Educational Research*. Vol. 1, No. 3, 283-299.

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988) Growth in reading and how children spend their time outside of school, *Reading Research Quarterly*, 23, 285–303.

Azeem M. (2013) *Development of Math Proficiency Test Based on Item Response Theory (IRT)*, Unpublished Ph.D. dissertation Division Of Education University Of Education, Lahore.

Benbow C.P. and Stanley, J. (1983) *Sex Differences in Mathematical Reasoning Ability, More Facts, Science,* 222: 1029-1031,http://www.vanderbilt.edu/Peabody/SMPY/ScienceMoreFacts.pdf

Bezzina, F.H (2010) Investigating gender differences in mathematics performance and in self-regulated learning: An empirical study from Malta, Equality, Diversity and Inclusion: *An International Journal*, Vol. 29 No: 7, pp.669 – 693.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999) *How people learn: Brain, mind, experience, and school*, (Washington, DC) National Academy Press. Available: http://books.nap.edu/catalog/6160.html. [July 10, 2001].

Brownell, W. A. (1935) *Psychological considerations in the learning and the teaching of arithmetic,* In W. D. Reeve (Ed.), The teaching of arithmetic (Tenth Yearbook of the National Council of Teachers of Mathematics, pp. 1–31) New York: Columbia University, Teachers College, Bureau of Publications.

Carpenter, T. P., Fennema, E., Fuson, K., Hiebert, J. Human, P., Murray, H., Olivier, A., & Wearne, D. (1999) *Learning basic number concepts and skills as problem solving In E. Fennema & T. A. Romberg, Mathematics classrooms that promote understanding,* (pp. 45–61). Mahwah, NJ: Erlbaum.

Carroll, J.B. (1996) *Mathematical abilities: Some results from factor analysis*, In R. J. Sterburg, & T. Ben_Zeev, (Eds.).The nature of mathematical thinking (pp.3-25), Mahwah, NJ: Lawrence Erlboum Assosiates.

Cohen, D. K., & Ball, D. L. (2000, April) Instructional innovation: Reconsidering the story. Paper presented at the meeting of the *American Educational Research Association*, New Orleans.

Davis, R. B., & Maher, C. A. (1997) *How students think: The role of representations*, In L. D. English (Ed.), Mathematical reasoning: Analogies, metaphors, and images (pp. 93–115). Mahwah, NJ: Erlbaum.

Donovan, M. S., Bransford, J. D., & Pellegrino, J. W. (Eds.). (1999) *How people learn: Bridging research and practice*, Washington, DC: National Academy Press. Available:

http://books.nap.edu/catalog/9457.html. [July 10, 2001].

Dossey, J. A., Mullis, I. V. S., Lindquest, M.M., & Chambers, D.L.(1988) *The mathematics report card: Are we measuring up?* Princeton NJ: Educational Testing Service

Dweck, C. (1986) *Motivational processes affecting learning,* American Psychologist, 41, 1040–1048.

English, L. D. (1997a) *Analogies, metaphors*, *and images: Vehicles for mathematical reasoning. In L. D. English (Ed.), Mathematical reasoning: Analogies, metaphors, and images* (pp. 3–18), Mahwah, NJ: Erlbaum.

Enreston, S.E.(1997) *Multi - component response models*, In van der Linden, W.J. & Hambleton, R. RK. (Eds.), Handbook of modern item response theory (pp.305-321). New York: Springer - Verlag.

Fennema, E., Carpenter, E. T., Jacob, V. R., Frank, M. L., & Levi, L. W. (1998) A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher*, 27 (5), 6-11.

Fuson, K. C. (1992b). Research on whole number addition and subtraction. In D. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 243–275). New York: Macmillan.

Geary, D. C. (1995) Reflections of evolution and culture in children's cognition, American Psychologist, 50(1), 24–37.

Gronlund, N. E. & Linn, R. L. (2005) *Measurement and assessment in Teaching*, New Dehli: Baba Barkha Nath Printers.

Hambleton, R. K. Swaminathan, H. & Rogers, H. J.(1991) Fundamentals of Item Response

Haladyna, T. M., & Downing, S. M. (1989) A taxonomy of multiple-choice item-writing rules, *Applied Measurement in Education*.

Halpern F.D; Benbow, C.P.;Geary, D.C,; Gur, R.C and Hyde, J S. (2007). *The science of Sex Differences in Science and Mathematics*, psychological science in the public interest vol. 8 no. 1 1-51

Hargreaves, M., Homer, M. & Swinnerton, B. (2008), A comparison of performance and attitudes in mathematics amongst the gifted. Are boy as better at mathematics or do they just think they are? Assessment in Education: Principles, Policy & Practice.

Hiebert and Carpenter, 1992, Hiebert and Wearne, 1996. For work in psychology, see Baddeley, 1976; Bruner, 1960, pp. 24–25; Druckman and Bjork, 1991, pp. 30–33; Hilgard, 1957;

Katona, 1940; Mayer, 1990; Wertheimer. (1959) Theory, London, *Sage Publications*.

Hiebert, J. (Ed.). 1986*) Conceptual and procedural knowledge,* The case of mathematics. Hillsdale, NJ: Erlbaum

Hiebert, J., & Wearne, D. (1996) *Instruction, understanding, and skill in multi digit addition and subtraction,* Cognition and Instruction, 14, 251–283.

Hyde, J. S., Fennema, E., & Lamon, S. (1990) Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139 –155.

Kilpatrick, J. (1985) Doing mathematics without understanding it, A commentary on Higbee and Kunihira. *Educational Psychologist*, 20(2), 65–68.

Lai, C.P. (1998) *Statistical modeling of student's performance in multi-component tasks*, Unpublished Ph.D. Thesis, Department of Learning and Educational Development, University of Melbourne

Li min, WU, M.(2003) *The Application of Item Response Theory to Measure Problem-solving Proficiencies,* Unpublished Ph.D dissertation. Department of Learning and Educational Development The University of Melbourne.

Maccoby, E. E., & Jacklin, C. N. (1987) Gender segregation in childhood, In H. Reese (Ed.), Advances in child development and behavior (Vol. 20, pp.239-288), New York: Academic Press.

Maher, C. A., & Martino, A. M. (1996) The development of the idea of mathematical proof: A 5-year case study, *Journal for Research in Mathematics Education*, 27, 194– 214.

Mayer, R.E. (1992) *Thinking, problem-solving, cognitive,* (2nd ed.) New York. Freeman and Company.

Mayer, R.E., & Hegarty, M.(1996) The process of understanding mathematical problems. In R. J. Sternberg & T. Ben_Zeev (Eds.), The nature of mathematical thinking (pp.29-53).Mahwah, NJ: Lawrence Erlboum Associates.

National Council of Teachers of Mathematics (1989) *Curriculum and evaluation standards for school mathematics*, Reston, VA: NCTM.

Mills, C. J., Ablard, K. E., & Stumpf, H. (1993) Gender differences in academically talented young students' mathematical reasoning: patterns across age and sub skills, *Journal Of Educational Psychology*, 85(2), 340-346.

Newman J. R. (Eds.). (1956). The *World of Mathematics,* New York: Simon and Schuster.

Pesek, D. D., & Kirshner, D. (2000) Interference of instrumental instruction in subsequent relational learning. *Journal for Research in Mathematics Education*, 31, 524–540.

Pólya, G. (1945) *How to solve it: A new aspect of mathematical method*, Princeton, NJ: Princeton University Press

Reckase, M.D.(1997) *A linear logistic multi-dimensional model for dichotomous item response data*, In W. J. Van der Linden & R. K. Hambleton (Eds.) *Handbook of modern item response theory* (pp.271-286).New York:

Resnick, L. B. (1987) *Education and learning to think,* (Washington, DC) National Academy Press.

Suen, H. K.(1990) *Principles of Test Theories,* Pennsylvania State University, Hove and London. Lawrence Erlbaum Associates, Publishers.

Schoenfeld, A. H. (1989) Explorations of students' mathematical beliefs and behavior, *Journal for Research in Mathematics Education*, 20, 33

Shulman, L. S. (1987) *Knowledge and teaching: Foundations of the new reform*, Harvard Educational Review, 57, 1–22.

Webb, N. L. (1997b, January) *Determining alignment of expectations and assessment in mathematics and science education,.*

Wright, B. D. & Linacre, J. M.(1994) Reasonable mean-square fit values. *Rasch Measurement Transaction,* 8:3 p.370 Retrieved from http://www.rasch.org/rmt/rmt83b.htm on 24.08.2013.

Yaffee, L. (1999) *Highlights of related research*, In D. Schifter, V. Bastable, & S. J. Russell with L. Yaffee, J. B. Lester, & S. Cohen, Number and operations: Making meaning for operations. Casebook (pp. 127–149). Parsippany, NJ: