

**BERND WIEGAND, KASSEL**

## **AN ANALYSIS OF REALITY-RELATED TIMSS-ITEMS**

### **Abstract**

The basis of the presented interpretations is a detailed analysis of TIMSS-Items taken from the Population 2-test. Analysis with similar results is also available for Population 3-Items (see WIEGAND 1999).

With this analysis of reality-related TIMSS-Items and results, interesting patterns become obvious which make it possible to develop a starting point for interpretations of the data. In this way one can identify specific strengths and weaknesses in German mathematics education, which is - in my opinion - one of the most interesting advantages of international comparative studies. Yet this potential has very seldom been used, since it is mainly aggregate results (mean scores or ranking lists) which have been discussed.

### **1. Basic concept of the analysis**

First of all I would like to describe and justify the selected methodological approach, which is distinctly different from the one in the German TIMSS report (BAUMERT ET AL. 1997).

- (1) Instead of using Rasch-scaled scores of the difficulties for single items, I consider the relative percentage of students with the correct answer. One reason for using percentages of correct responses is that the results for each item and each country are given in this way (<http://timss.bc.edu/timss1995i/Database.html>) and not as Rasch-scaled scores. In addition, percentage correct scores are easier to interpret than the Rasch-scaled scores, indeed there is no advantage in using Rasch-scaled scores for the following analysis.
- (2) One main reason why a detailed analysis like the one presented here is necessary and helpful is that the consideration of mean scores over all items is superficial since it is not possible to identify specific strengths and weaknesses of the mathematical education of a country. There are a variety of methodological criticisms of international comparison studies in this respect. For example, MCKNIGHT/VALVERDE (1999, p. 59) stated that in studies such as TIMSS it is generally not possible to consider curricular differences between the participating countries because of the large number of nations. Therefore, overall results are restricted in their value:

“ ... globally-scaled scores based on high-level aggregations of very disparate items at best measure only very general mathematics skills, and are thus essentially insensitive to curricular differences. ... In the meantime, it might be worth taking a closer look at content scores to see if, even at what is still a fairly global level, decreased aggregation can yield more curricularly related results.” (1999, p. 59)

WOLFE (1999, pp. 225) confirmed this criticism with a detailed analysis of the SIMS-data. He argued that in great international comparative studies there exists a strong dependency between the results of the study and the selection of the items used in the test. According to his analysis, the number of items of a certain kind determines the overall result, hence the result is ultimately determined by the curricular emphasis of the country's educational system. He found a strong dependency between the variable “opportunity to learn” and the test-score by analysing SIMS-data of different countries:

“For example, if the test design or weighting favours the geometry topic, then countries with relatively high geometry performance will rank higher overall - an example of international measurement bias.” (1999, p. 228)

Statistically this can lead to a measurement error:

“Unless there is an exactly parallel structure of knowledge, learning and language in different student populations, some items will be relatively harder in some populations than in others. The random (unexplained) variation within a content sub-domain for which a score is to be developed must be regarded as an error because, once we have formed the score, we are not interested in the particular set of items that were used ... .

For example, if we have a small number of items in a topic, and the items vary differently in difficulty from country to country, the overall ranking of the countries in the topic will have lots of international measurement error.” (1999, p. 228)

This criticism by WOLFE is confirmed by RAMSEIER 1997, who has analysed TIMSS (Population 2). He discovered that the rank order concerning the results in natural sciences of various countries alters distinctly when, for instance, only items with low cognitive requirements and high terminological difficulty are taken as a basis. For example, the ranking of Switzerland in TIMSS (Population 2, Science) is worse if only items of this kind are considered while their position is far better if only items with high cognitive requirements and low terminological difficulty are considered (see Ramseier 1997, p. 43).

Consequently, there is a distinct need for a detailed consideration of the analysis and the results of single items as in the presented work.

My starting point for an interpretation is the analysis of items in classes of items (here: reality-related items). The results of these items are considered and interpreted by elaborating patterns in the classes of items.

- (3) Isolated percentages of correct responses from German students (e.g.: 45 % of the German students answered item I 2 correctly) are not interpretable without a relative value. I have therefore compared the percentage of correct responses from German students with that from all students in all participating countries. This is a relative value which is, in my opinion, independent of specific characteristics of each country. It is also easy to interpret, since the overall percentage of correct responses is almost identical with the overall percentage of the German students:- in grade 8, a mean of 54 % over all items in Germany, and a mean of 55 % over all countries: in grade 7, an identical mean of 49 % in Germany and over all countries. Therefore items in which the German students performed much better mark their (relative) strengths, while items in which the German students performed much worse mark their (relative) weaknesses.

I would like to distinguish roughly between two kinds of items:

- mathematical items, for example item M 8: “Multiply:  $0.203 \times 0.56 =$ ”
- reality-related items, for example N 16:

“Jan had a bag of marbles. She gave half of them to James and then a third of the marbles still in the bag to Pat. She then had 6 marbles left. How many marbles were in the bag to start with? “ (multiple choice item, alternatives: 18; 24; 30; 36)

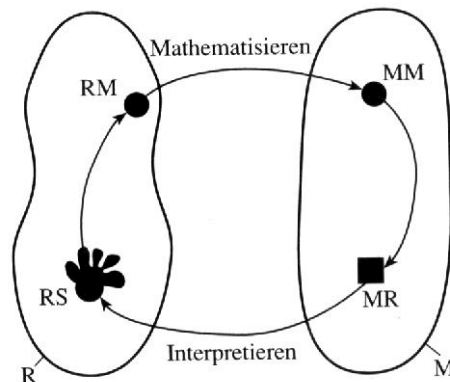
The following analysis is primarily based on reality-related items.

Section 2 describes what is meant by reality-related items. For the solution of these kinds of items Grundvorstellungen (one cannot translate this term exactly, “concept image” or “ground rules” means almost the same) of the mathematical content are necessary (Section 3). Then (Section 4) some arithmetic- and algebra-items will be analysed and the results will be interpreted (Section 5). In Section 6 some other items will be analysed and the results will be interpreted. Finally, in Section 7 some obvious and sensible modifications to German mathematics instruction will be suggested.

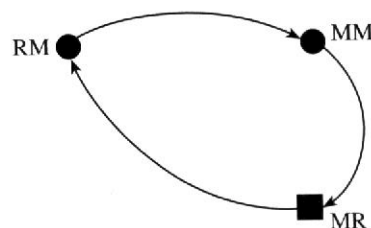
## 2. Reality-related Items

The following description in Sections 2, 3 and 4 is similar to the one in BLUM/WIEGAND 1998. Reality-related items are - in my terminology - items in which the mathematical content is in any way related to “reality”.

One can distinguish - given the complexity of such items - between problem-solving (in the sense of the modelling process, see Figure 1), the use of standard models and word problems<sup>1</sup>. (see BLUM 1996).



Word problems are very relevant for mathematics instruction in school. From the outset, the real situation (RS) is structured in such a way that a real model (RM) nearly exists. The process of solving such an item then consists of five steps (see Figure 2, WIEGAND 2000, p. 39): From the text given in the item (AT) to the kernel of the text (Kern AT) to the real model (RM) and then to the mathematical model (MM), mathematical results (MR) and back to the situation given in the text.



<sup>1</sup> In my terminology word problems are mathematical problems presented as a task in the form of a short text.

Reality-related TIMSS-Items are mainly word problems (because of the time limitation in such a test). The difficulties for students in solving these items can lie - according to many investigations - in all three steps:

- the mathematisation, i.e. the efficient reading of the text of the item and the forming of an equation, a function, a geometrical figure, etc. In Item N 16 for example one has to translate operations with marbles into mathematical operations with numbers;
- the mathematical working process. In N 16 for example one needs to have knowledge about fractions;
- the interpretation, i.e. the transference of the mathematical results back to the real situation. In N 16 for example one has to interpret the result as the number of marbles in the initial situation.

The third step— as in the “marble-item” N 16 - is much easier in many word problems than the first two steps, which are the main problem areas for students in solving such an item, especially the translation from reality into mathematics. A student can find a suitable approach to the problem if he has an adequate knowledge of the inherent mathematical terms and operations and if he can activate these Grundvorstellungen (basic concepts) in this particular context. This is why these Grundvorstellungen (GV) play an important role in the process of solving word problems.

### 3. Grundvorstellungen

The underlying theory of Grundvorstellungen is outlined below (see VOM HOFE 1995).

The question of adequate concepts of mathematical terms and operations has a long tradition, beginning with Pestalozzi’s “Anschauungen”. The term “Grundvorstellung” was first used by Oehl and was then developed further by Griesel.

In his theory, vom Hofe defined the term “Grundvorstellung” as the relationships between the mathematical content, the real context and the formation of a concept by the individual. Grundvorstellungen mediate especially between mathematics and reality and they can help in solving a word problem by making the mathematisation step easier.

There are usually several, quite often even diverse Grundvorstellungen for each mathematical content; here are three examples, which will be used in Sections 4 and 5.

1) I distinguish between three basic Grundvorstellungen for fractions (always with the example  $\frac{2}{3}$  compare with PADBERG 1995):

- ◆ fractions as rates:  $\frac{2}{3}$  as two thirds of a pizza or two pizzas for three persons;
- ◆ fractions as operators:  $\frac{2}{3}$  of a certain amount, for example 60 persons (see item N 16);
- ◆ fractions as ratios:  $\frac{2}{3}$  as a ratio, for example two parts of apple-juice and three parts of mineral water.

2) One can distinguish (see PADBERG 1996) between four Grundvorstellungen for subtraction (always with the example 8-5):

- ◆ subtraction as taking away: One child has eight marbles and gives away five marbles, how many marbles are left? (see item N 16)
- ◆ subtraction as comparing: One child has eight marbles, another child has five marbles, how many more marbles has the first child than the second?

- ◆ subtraction as adding to something: One child has five marbles, how many marbles does it need to have eight marbles?
- ◆ subtraction as combining: One child has eight marbles, coloured red and blue. They have five red marbles, how many blue marbles do they have?

3) Like in MALLE (1993) one can distinguish between three “Grundvorstellungen” of the term variable. They are also relevant for terms and equations:

- ◆ variable as an unknown, a number not further specified (Gegenstands-Grundvorstellung);
- ◆ variable as a fill-in for a number (Einsetzungs-Grundvorstellung);
- ◆ variable as a symbol, which one can manipulate according to certain rules (Kalkül-Grundvorstellung)

In item N 16 (“marble-item”) the first Grundvorstellung will be activated. Once a term or equation has been set up, it can be manipulated according to certain rules. Since the task is a multiple choice item, the Einsetzungs-Grundvorstellung can also lead to a solution by using the given alternatives. Generally, a systematic trial can lead quickly to a solution. This can be achieved, without forming a term or an equation, by checking the alternatives against the text of the item.

#### 4. Analysis of reality-related TIMSS-Items

In this Section, five selected reality-related TIMSS-Items from the domain of arithmetic and algebra will be analysed with regard to the following aspects:

- A) involved “Grundvorstellungen” (type, quantity and their interaction; relevant for all of the three steps mentioned in 2. to solve a word problem);
- B) process of mathematisation (the first step in solving a word problem; familiarity with the reality-based context is one of the important factors);
- C) process of solving the problem mathematically (the second step in solving a word problem; one factor is the quality and quantity of the necessary elements, another is the degree of difficulty of the numbers used).

Of course, these three aspects are related to each other, especially A) to B).

4.1 The “marble-item” (N 16) has already been analysed in Sections 2 and 3. To summarize:

- A) Required are
  - the operator-Grundvorstellung of fractions,
  - one or two Grundvorstellungen (according to the process of solution) of variables,
  - the taking-away-Grundvorstellung of subtraction to solve the item. There are several different concepts involved, each of which are related;
- B) The context is familiar, but the mathematisation is not easy because of the complexity of the real situation;
- C) The calculation consists of several steps. Each individual step is easy, but all of them have to be solved to obtain the final result.

4.2 The “bus-item” (I 2)

“Two groups of tourists each have 60 people. If  $\frac{3}{4}$  of the first group and  $\frac{2}{3}$  of the second group board buses to travel to a museum, how many more people in the first group board buses than in the second group?” (multiple choice item, alternatives: 2, 4, 5, 40, 45)

A) Required are:

- the operator-Grundvorstellung of fractions;
- the comparing-Grundvorstellung of subtraction (more difficult than the others, especially in the formulation “how many more ... than ...”) to solve the problem.

B) The context is familiar. However the translation into a mathematical model is only possible if one understands the real situation and is able to combine the concepts mentioned in A).

C) The numbers used are easy, but the calculation consists of several steps.

#### 4.3 The “ratio-item” (M 6)

“A class has 28 students. The ratio of girls to boys is 4 : 3. How many girls are in the class?”

A) Required are:

- the ratio-Grundvorstellung of fractions and - depending on which way is chosen to solve the problem either
  - the operator-Grundvorstellung or
  - a concept of proportionality.

The ratio-Grundvorstellung is not usual in German mathematics education in grade 7 and 8.

B) Again, the context is familiar, but the mathematisation requires the student to know what ratio means.

C) The numbers used and the calculations are easy.

#### 4.4 The “cake-item” (P 14)

“Janis, Maija and their mother were eating a cake. Janis ate  $\frac{1}{2}$  of the cake. Maija ate  $\frac{1}{4}$  of the cake. Their mother ate  $\frac{1}{4}$  of the cake. How much of the cake is left?”

(multiple choice item, alternatives:  $\frac{3}{4}$ ;  $\frac{1}{2}$ ;  $\frac{1}{4}$ ; none)

A) Required are:

- the rates-Grundvorstellung of fractions;
- the taking-away-Grundvorstellung of subtraction to solve the problem; both are less demanding.

B) The context is familiar and the mathematisation is very easy, if the text of the item is followed step by step.

C) The numbers used in the calculation are easy.

#### 4.5 The “hat-item” (Q 1)

“Juan has 5 fewer hats than Maria, and Clarissa has three times as many hats as Juan. If Maria has  $n$  hats, which of these represents the number of hats that Clarissa has?”

(multiple choice item, alternatives:  $5 - 3n$  ;  $3n$  ;  $n - 5$  ;  $3n - 5$  ;  $3(n - 5)$  )

A) One needs to have

- the Gegenstands-Grundvorstellung of variables
- the taking-away-Grundvorstellung of subtraction (in the formulation “fewer than”)
- the operator-Grundvorstellung of multiplication to solve the problem.

B) The context is unusual, but easy to understand. For the mathematisation one needs to structure the text, especially because of the succeeding if-clause.

C) There is nothing to do.

## 5. Interpretation of the results

The following results are based upon the TIMSS-2 data available on the Internet.

Looking at the results of reality-related items you not only find interesting patterns of the items presented in Section 4, but of the whole class of items.

The results (percentage of correct responses in grade 7 and grade 8 for Germany and all participating countries together; the results of the analysis in Section 4) of the reality-related items of Section 4 are shown in Figure 3. Items in which the German students performed much better mark their (relative) strengths, while items in which the German students performed much worse mark their (relative) weaknesses.

<b><u>Figure 3:</u></b> <b><u>reality-related items</u></b> <b><u>(arithmetic)</u></b>	percentage of correct responses in Germany	percentage of correct responses in all participating countries	
I 2: “bus-item” (grade 7) (grade 8)	45 % 49 %	52 % 57 %	operator-GV and comparing-GV, difficult mathematisation
M 6: “ratio-item” (grade 7) (grade 8)	19 % 30 %	30 % 37 %	ratio-GV and operator-GV
P 14: “cake-item” (grade 7) (grade 8)	83 % 77 %	72 % 76 %	part-GV and taking-away-GV, direct translation

<b><u>Figure 4:</u></b> <b><u>reality-related items</u></b> <b><u>(arithmetic)</u></b>	percentage of correct responses in Germany	percentage of correct responses in all participating countries	
Q1: “hat-item” (grade 7) (grade 8)	27 % 41 %	37 % 47 %	difficult mathematisation Gegenstands-GV
N 16: “marble-item” (grade 7) (grade 8)	37 % 35 %	43 % 48 %	Gegenstands- or Einsetzungs-GV, operator-GV

Thus the following patterns can be identified (see BLUM/WIEGAND 1998, p. 32 and WIEGAND 1998, p. 21):

If the word problem - as in P 14 - requires only some simple Grundvorstellungen and the mathematisation is directly possible, the results of German students are (relatively) good.

But if

- more demanding or less familiar Grundvorstellungen are required - as in the “bus-item” or the “ratio-item” - or
  - several Grundvorstellungen have to be combined - as in the “marble-item” - or
  - the mathematisation is difficult - as in the “hat-item” - or
  - if more than one of these aspects coincide,
- then the results of German students are relatively worse than the international mean.

## 6. Interpretation of the results of other items

If you look at the results of non-reality-related arithmetic items, you can find the following trend (see WIEGAND 1998, p. 21 and Figure 5): while German students usually perform better in items dealing with fractions in grade 7, they are less successful in grade 8. In some cases they even fall behind the international mean.

<b><u>Figure 5: non-reality-related items (fractions)</u></b>		percentage of correct responses in Germany	percentage of correct responses in all participating countries
J 12: “Divide: $\frac{8}{35} : \frac{4}{15}$ ”	grade 7	46 %	36 %
	grade 8	44 %	43 %
P 16: “Write 0.28 as a fraction reduced to its lowest terms”	grade 7	37 %	30 %
	grade 8	33 %	34 %

If you look at the results of non-reality-related algebra items (which do not include any modelling), the trend is not clear, but there are some items with better results for German students (see Figure 6). Item Q 2 does not seem to require simple manipulation, but a more advanced knowledge about variables, combined with a Grundvorstellung of fractions.

The trend described above can also be found in other areas. For example, the slightly better performance of German students in the area “data representation, analysis and probability” (BEATON ET AL. 1996, p. 41/42) can be explained by the small number of items which require a difficult mathematisation or less familiar Grundvorstellungen. Two examples of typical items from this content area and their results can be found in Figure 7 (see also BEATON ET AL. 1996, p. 80-86).



<b><u>Figure 7: two items out of the content area “data representation, analysis and probability”</u></b>		percentage of correct responses in Germany	percentage of correct responses in all participating countries
P 17: “This table shows temperatures at various times during the week. ... Which thermometer shows the temperature at 8 p.m. on Monday?” (four alternatives)	grade 7 grade 8	94 % 94 %	79 % 82 %
O 1: “The graph shows the distance travelled before coming to a stop after the brakes are applied for a typical car travelling at different speeds. ... A car travelling on a highway stopped 30 m after the brakes were applied. About how fast was the car travelling?” (four alternatives)	grade 7 grade 8	68 % 69 %	51 % 58 %

<b><u>Figure 6: non-reality-related items (algebra)</u></b>		percentage of correct responses in Germany	percentage of correct responses in all participating countries
P 15: “Which of these expressions is equivalent to $y^3$ ?” (alternatives: $y+y+y$ ; $y \cdot y \cdot y$ ; $3y$ ; $y^2+y$ )	grade 7 grade 8	60 % 73 %	55 % 66 %
Q 2: “Subtract: $\frac{2x}{9} - \frac{x}{9} =$ ” (alternatives: $\frac{1}{9}$ ; $2$ ; $x$ ; $\frac{x}{9}$ ; $\frac{x}{81}$ )	grade 7 grade 8	33 % 38 %	41 % 51 %

In addition to this, German students generally perform worse than the international mean in items in which the term “ratio” is used or which require the ratio-Grundvorstellung (as in item M6). One example for this is shown in Figure 8.

<b><u>Figure 8: one item for the ratio- Grundvorstellung</u></b>		percentage of correct responses in Germany	percentage of correct responses in all participating countries
P 8: “What is the ratio of the length of a side of a square to its perimeter?” (alternatives: $\frac{1}{1}; \frac{1}{2}; \frac{1}{3}; \frac{1}{4}$ )	grade 7	37 %	49 %
	grade 8	45 %	56 %

## 7. Possible reasons and conclusions

There are various reasons for the poor results of German students e.g. the low status of mathematics and science education, lack of eagerness to learn, etc. Mathematics education is not the only area to which these comments apply and such failings can only be changed in the long term. However there are several other reasons, which are, in my opinion, related to mathematics instruction in Germany.

Possible criteria for identifying such reasons are qualitative-related studies like the TIMSS-Video-study (see the cultural scripts of mathematics instruction in Japan, the USA and Germany presented in BAUMERT ET AL. 1997) or the classroom observations of German and English mathematics instruction as a part of the Kassel-Exeter study (see the idealised types of mathematics teaching in both countries presented in KAISER 1999a and KAISER 1999b).

Given this background it seems reasonable to suggest that German mathematics instruction is (too) strongly related towards routines and standard tasks. Students also tend to forget much of the content they have learned for the last test (see the comment of the association of DMV, GDM and MNU on TIMSS). Therefore the demands for less standard tasks and more problem solving, for more reality-orientation, connected learning and intensified repetition is supported by the analysis presented above.

One central point which, in my opinion, makes empirical studies like TIMSS valuable is that they can help to identify specific strengths and weaknesses in the mathematics and science education in each country.

A project that is starting to change mathematics and science instruction in Germany in the way described above is the BLK-Modellversuchsprogramm (see e.g. HENN 1999). Even with this programme, however, I doubt that the German students can become as good as the Japanese students of the same age, because of the totally different cultural and social backgrounds. We should look primarily at our European neighbours (like the Netherlands or Switzerland), which have also achieved excellent results. I therefore await the results of the planned video-study in Switzerland with great interest.

## References

- BAUMERT, J. et al. (1997): TIMSS - Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Opladen: Leske und Budrich
- BEATON, A.E. et al. (1996): Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study. Chesnut Hill: Boston College
- BLUM, W. (1996): Anwendungsbezüge im Mathematikunterricht - Trends und Perspektiven. In: G. Kadunz et al. (Ed.): Trends und Perspektiven. Beiträge zum 7. Internationalen Symposium zur "Didaktik der Mathematik" in Klagenfurt vom 26.-30.9.1994, Wien: Hölder-Pichler-Tempinsky, p. 15-38
- BLUM, W.; NEUBRAND, M. (Hg.) (1998): TIMSS und der Mathematikunterricht. Hannover: Schroedel
- BLUM, W.; WIEGAND, B. (1998): Wie kommen die deutschen TIMSS-Ergebnisse zustande? Ein Interpretationsansatz auf der Basis stoffdidaktischer Analysen. In: Blum, W.; Neubrand, M. (Hg.): TIMSS und der Mathematikunterricht. Hannover: Schroedel, p. 28-34
- HENN, H.-W. (1999): Das BLK-Projekt - Herausforderung und Chance. In: Henn, H.-W. (Hg.): Mathematikunterricht im Aufbruch. Hannover: Schroedel, p. 7-13
- KAISER, G. (1999a): Unterrichtswirklichkeit in England und Deutschland. Vergleichende Untersuchungen am Beispiel des Mathematikunterrichts. Weinheim: Deutscher Studien Verlag
- KAISER, G. (1999b): Comparative Studies on Teaching Mathematics in England and Germany. In: Kaiser, G.; Luna, E.; Huntley, I. (Eds.): International Comparisons in Mathematics Education. London: Falmer Press, p. 140-150
- MALLE, G.(1993): Didaktische Probleme der elementaren Algebra. Braunschweig: Vieweg
- MATHEMATIK LEHREN 90 (1998): TIMSS - Anstöße für den Mathematikunterricht (Themenheft). Seelze: Friedrich Verlag
- MCKNIGHT, C.C.; VALVERDE, G.A. (1999): Explaining TIMSS Mathematics Achievement: A Preliminary Survey. In: Kaiser, G.; Luna, E.; Huntley, I. (Eds.): International Comparisons in Mathematics Education. London: Falmer Press, p. 48-67
- OEHL, W. (1970): Der Rechenunterricht in der Hauptschule. Hannover: Schroedel
- PADBERG, F. (1996): Didaktik der Arithmetik. Heidelberg: Spektrum
- PADBERG, F. (1995): Didaktik der Bruchrechnung. Heidelberg: Spektrum
- RAMSEIER, E. (1997): Naturwissenschaftliche Leistungen in der Schweiz. Vertiefende Analyse der nationalen Ergebnisse in TIMSS. Bern: Amt für Bildungsforschung
- VOM HOFE, R. (1995): Grundvorstellungen mathematischer Inhalte. Spektrum: Heidelberg
- WIEGAND, B. (1998): Stoffdidaktische Analysen von TIMSS-Aufgaben. In: mathematik lehren 90 (1998), p. 18-22
- WIEGAND, B. (1999): TIMSS als Spiegel für Defizite im deutschen Mathematikunterricht der Sek. II - Analysen von Aufgaben aus TIMSS-3 und Interpretationen der Ergebnisse. In: Beiträge zum Mathematikunterricht 1999. Hildesheim: Franzbecker 1999, p. 594-597
- WIEGAND, B. (2000): Mathematische Anwendungsfähigkeiten. Detailanalysen von TIMSS und Kassel-Exeter-Studie. Hildesheim: Franzbecker 2000
- WOLFE, R.G. (1999): Measurement Obstacles to International Comparisons and the Need for Regional Design and Analysis in Mathematics Surveys. In: Kaiser, G.; Luna, E.; Huntley,

I. (Eds.): International Comparisons in Mathematics Education. London: Falmer Press, p. 225-240

Bernd Wiegand  
Universität Gesamthochschule Kassel  
Fachbereich 17-Mathematik/Informatik  
Heinrich-Plett-Str. 40  
34132 Kassel