Peter Krope, Wilhelm Wolze

Ordinary Language and Scientific Language. Test Validity from a Constructivistic Point of View

The Problem

The authors of this contribution have developed with their research team a course in the area of elementary mechanics. The course is being tested in three countries, Germany, Poland and Hungary. The experiment serves to test the educational objectives. The didactics especially designed for the course is supposed to lead to learning results which – in the sense of the taxonomy of Bloom et al. (1973) – go beyond only reproducing knowledge. The didactics is also supposed to be applicable internationally.

This contribution serves as methodical preparation for the international study. According to Postlethwhaite (1988: XVII) its subject is the method of the comparative education. That is the comparative method.

The descriptions of the comparative method vary. In essence its task is to describe relations as equal and unequal. The reference level for the comparison is the language and not what is talked about (Krope, Wolze, 2003). The language as a basis for comparisons will be analysed in the following. The purpose is to determine the importance of expressions used in the ordinary language and in the scientific language in order to ensure the validity of the test results.

Scientific and Non-scientific Language

The study prepared in this contribution is based on the Constructivism that has been found by Kamlah, Lorenzen (1973) and Lorenzen (1987) and has been further developed by Janich (2001). To differentiate from Radical Constructivism the position will be called 'Methodical Constructivism'. An encyclopaedic description is given

by Mittelstraß (1980, 1984, 1995, 1996). The Methodical Constructivism has been chosen as an epistemological position which involves the attempt to introduce science in understandable, circle-free, and non-dogmatic procedures.

The comparative method of the study is distinctively based upon the scientific language of the Methodical Constructivism. This language is called 'ortholanguage'.

To differentiate from the term 'ortholanguage' the term 'living world' (*Lebenswelt*) is used. This is an expression that above all the philosopher Edmund Husserl (1976) has made well known. With the term 'living world' the area of the pre-scientific experiences can be differentiated from the scientific experiences. As a criterion for differentiation Böhme (1979: 124) works out, based on an analysis of Merleau-Ponty (1966), the specific relation of perception and organism:

> Life worldly experience is essentially sensual experience; it is bound to the human body.

In contrast to this Böhme characterises the scientific perception as the registration of interaction processes of the system. Life worldly knowledge serves the immediate experience in a complex environment. It is – in contrast to scientific knowledge – situation related and therefore context open. The transition from the life worldly to the scientific sector is based on a change which transits the essential qualities of the life worldly experiences. Bachelard (1978) characterises the transition from the life worldly perception to the scientific perception by means of discontinuities of epistemological break ups so that the scientific perception is not an immediate continuation of the life worldly perception. Both types of perception have their own function. The questions are: Which characteristics has the life worldly perception? What is the relation between the scientific perception and the life worldly perception? How changes or should change the life worldly perception in connection with the scientific perception?

In this contribution the term 'life worldly' is used as a synonym for 'pre-scientific'. The classification of the language into two categories is done for the purpose of the following considerations. It is not expected that in empirical analyses they go beyond the purposes of the contribution proving to be selective and exhaustive.

The common language in particular often leads to misunderstandings because of its context openness. In the scientific language terminological systems are designed to avoid misunderstanding. For that purpose explicit terms are being introduced and rules for its use are agreed upon. For example pupils watching a film about life in the water in their biology class use the common language expression 'whale-fish' when talking about their first observations. After the successful conclusion of the biology class they start a differentiation of fish. Whales reproduce like mammals. From now on the pupils avoid the common language expression 'whale-fish' and define in a technical term a standardised use. For the observed tooth whales they introduce the term 'Odontoceti' according to the international rules of zoological definition. Due to the terminological agreement as in the example a scientific language is build up in such a way that pre-scientific and scientific descriptions of circumstances differ in

two ways. First of all a difference with regard to the contents can be found. In the example in the beginning of the class the pupils count the whale pre-scientifically as a fish and after the conclusion of the class in the scientific sense to the mammals. Secondly there is a formal difference when in the example it is talked in the common language about 'whales' and in the scientific language about 'Odontoceti'. This differentiation in contents and in forms is to be found in the Methodical Constructivism (compare Kamlah, Lorenzen 1973: 70–116) as well as in traditional scientific positions (compare Körner, 1980: 727–728).

Which consequences has the differentiation? In sports and in play it has been so far observed how a ball is thrown against a wall and rebounds again. If a physicist is asked for the answer why the ball rebounds he will in an extreme case answer with a formula whereas a pupil describes his daily knowledge in a longer and wordy explanation (compare table 1).

Table 1.
Common language and scientific answers to the question why a ball rebounds after it has been thrown vertical to a wall

Answers to the question why a ball rebounds from a wall	
In common language	**In scientific language**
'That is a physical law. After rebounding of the wall the ball at first does not know where to go. But because it wants its old shape back it pushes itself from the wall and therefore rebounds. The thrower knows the law and uses it to be able to catch the ball again.'	$F(K_1; K_2)$

The common language answer on the left side comes from a pupil of a secondary school in Gdańsk. The formula on the right side specifies the definition: The reason for changes in movement of a body K_1 is the exerted strength F of another body K_2 towards K_1 "$F(K_1; K_2)$".

The Achievement Test

The differentiation in contents and in forms between scientific and pre-scientific language raises a problem for the teaching and the learning. The problem is that learners usually cannot overcome the outlined gap. Pupils that are introduced to a science are in general more likely to be able to reproduce the scientific knowledge by for example quoting formulas. Skills that go beyond reproducing knowledge can be proved a lot less. To analyse this problem and its possibility for solutions the authors developed the course about elementary mechanics.

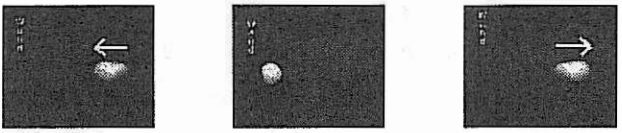
The check-up of the course in an international comparison is carried out with a test. To assure the objectivity the empirical analysis is being carried out with test items of the multiple-choice-type.

The mechanics test designed for the analysis contains altogether 13 items. Each item has only one right solution which is to be chosen from 4 alternatives. In table 2 a test item is described together with the instruction. The right answer is marked (b). The elements of a multiple-choice item are the instruction in the head, the illustration as well as the answers. The wrong alternatives a, c and d in table 2 are supposed to be so attractive that they distract from the right answer. The distraction function gave the wrong alternatives the name 'distracters'.

Table 2.
An item of the mechanics test

Instruction
Following you will find an item. In the item you are being asked a question. Below the question are four answers. Which answer from the physical point of view is the most applicable one? Please only cross one answer!

Item
A ball is thrown vertically to a wall. It rebounds. Why does it rebound?



a) The ball wants its old shape back.
b) The wall exerts strength towards the ball.
c) The ball exerts strength towards the wall.
d) The ball pushes itself off the wall.

The correct answer is marked (b)

The function of a test item to distract from the right answer will be explained with an example. During the development of a test to check up the Latin vocabulary it was expected of students to find for the Latin word 'studium' in an open answer a good German equivalence. The answer of 92% of the students was 'the study'. The final multiple-choice item contained as solution the term 'the engaged activity'. Besides this change the introduction of an attractive alternative ('the intellectual capability') made the exercise so difficult that at the survey eventually only 79% of the students selected the correct solution (compare table 3).

Table 3.
A multiple-choice item from a Latin test according to Dominick, Krope (1972: 80)

Studium
a) The university (5)
b) The intellectual capability (12)
c) The stupidity (1)
d) The **engaged activity** (79)

In the brackets the frequency is given with which the alternatives in the main survey of the study were chosen. 3% of the tests persons skipped the task. The correct solution is marked in the table (d)

The mechanics test shows a special feature. Firstly, in every item two alternatives are ortholinguistically formulated with the help of the terminology of the elementary mechanics. In table 2 that are (b) and (c). Secondly, two alternatives in each item are formulated in the everyday language. To construct the wrong common language expressed alternatives two procedures could be chosen from. With the textbook-method the wrong alternatives have been designed according to the specifications of textbooks of the physics-didactics. With the living world-method pupils have been asked to answer the questions in an open wording. The most frequently given open wordings were used for the development of the wrong alternatives. In the alternatives (a) and (b) in table 2 a part of the above quoted open answers of the pupil from Gdańsk can be recognised.

Living World and Everyday Language

Various factors can determine the kind of pre-scientific statements made by a person. To these factors could belong the gender, the age and the place of living. Schoolboys might make other pre-scientific remarks than schoolgirls, young people different ones than older people and people living in Poland again different ones than people in Germany. They belong to a different living world which might be linked to the different habits of the common language. The differences can be so great that young people from one district of a city can not understand the ones from another district when they use their common language. In the pre-scientific presentation of circumstances we are usually more likely to chose the language of our own rather than of a foreign living world.

From the explanations about the construction of the test and about the living world two assumptions can be concluded. Assumption number 1 says: If the wrong test answers come from the living world of a person these are more attractive and more likely to be chosen than as if they are not from their living world.

This assumption has been tested at a school in Gdańsk with two versions of the mechanics test: with a living world-version and a textbook-version. For the living world-version the wrong answers were developed from the wrong open answers of the pupils of Gdańsk. The textbook-version was first designed from didactics-books of the physics. The ortholinguistic answers of which one contained the right solution were in both versions identical. Both versions are available in a Polish translation. Half of the pupils of the school where the common language formulation comes from worked with the life worldly-version of Gdańsk of the test. The other half was given the textbook-version of the test. The allocation was done at random.

The assumption number 2 is: If a test person comes from the same living world where the common language phrased distracters derive from the distracters make it difficult to give a correct answer. The appropriate hypothesis says: The item with life worldly distracters is difficult, the right answer is less frequently marked.

Table 4:
Test results of the mechanics test from experiment one, survey in Gdańsk

	Test	
	Life worldly version	Text book-version
$\bar{x}$	4.11	5.89
N	18.00	18.00
s	2.03	2.89

The test persons come from the living world from which the life worldly distracters have been phrased. Explanation: $\bar{x}$ – arithmetic mean, N – number of test persons, s – standard deviation

The results are shown in table 4. As expected the life worldly test version is difficult. The pupils prefer the linguistic behaviour of their living world and are therefore being distracted from the correct answers. The differences are significant with a error probability of 2.5%.

In a second experiment the opposite assumptions has been analysed. The hypothesis is: If a test person comes from a different living world as to where the life worldly distracters derive from it is easier to give the correct answer.

In the second experiment at a secondary school in Kiel again the wrong open answers of the pupils have been gathered. Also with these wrong answers a life worldly-version of the test has been designed. The test persons in this second experiment were not the pupils but older people of the population of Kiel. Namely 75 older people from Kiel got the test with the at the school developed common linguistically wrong formulated alternatives while another 75 older test persons from Kiel were given the text book-version. The allocation was likewise done at random. Also in this second experiment again the ortholinguistically formulated answers of which one contained the right solution were identical in both versions.

The results of the second experiment are shown in table 5. Also in this experiment the assumption about the connection between the living world and the test results has been confirmed. The common linguistic explanations of an unfamiliar living world were more easily rejected. Because of that the correct answers became more attractive. Over all the life worldly-version of the test has become easier. There were more good results. The differences are significant with a error probability of about 15%.

This second experiment was carried out by students of the university within the population of Kiel. The students were asked to keep a record of the reactions of the test persons in regard to the test items. From the records it became evident that the older people who were tested and the younger pupils from whom the wrong answers were derived indeed belonged to different living worlds. The students had to listen over and over again how silly the wrong answers in the test have been formulated and how amateurish the professors of the university of Kiel dealt with physics knowledge.

Table 5.
The test results of the mechanics test of experiment 2, survey in Kiel

	Test	
	Life worldly version	Text book-version
$\bar{x}$	4.20	3.90
N	75.00	75.00
s	1.95	1.60

The test persons do not come from the living world from which the live worldly distracters have been formulated. For explanations see table 4

Consequences for the Everyday Language and the Scientific Language

The results of the empirical analysis which was carried out in Gdańsk and Kiel are temporary findings. The effects in particular of the small number of test persons are to be clarified at further tests. It is to be expected that the results of the experiment conducted in Kiel would have turned out more clearly if the group of test persons would have been less heterogeneous.

The following is about the question which consequences are to be drawn if the results are confirmed in further experiments. To be discussed are the functions of common and scientific language.

In essence the task of a comparative study is to describe relations as equal and unequal. For example two people might be asked why a ball that has vertically been thrown to a wall rebounds again. Both people might answer by marking alternative b in the test item of table 2: 'The wall exerts strength towards the ball'. According to the results of the study conducted in Kiel and Gdańsk it can not be concluded with certainty that both people show the same physical abilities with their answers.

What is the reason for the insecurity? Sentences which describe symmetric, transitive and reflexive relations in a defined area are called 'equal'. Equality is given through the relation of equivalence. The equalities which stand in the foreground in comparative empirical studies are the intentional and the extensional equivalence. The properties expressed by a predicator are called 'intention' (in the colloquial language: Content). The class of objects which show this property are called 'extension' (in the colloquial language: Extent). In the example both people marked alternative b. According to the presented empirical study the physical abilities are different to the highest degree if the persons and the common linguistic distracters derive from different living worlds. The test result refers to extensional equivalence. The test misses its aim, to make statements about intentional equivalence, as long as due to disregard of the living worlds the questioned properties stay indefinite.

What are the consequences? The beginning with life worldly expressions is a prerequisite of all comparative empirical studies and – since life worldly expressions often lead to misunderstandings – the use of a terminology is another. Both prerequisites are given with the scientific language of the Methodical Constructivism, the ortholanguage. The constructive build-up of the language starts on the pragmatic level. The beginning of the build-up of the language is being transferred into the living world. Applied are terms of the common language. From these the building-up proceeds via a semantic level to the syntax of an interpreted formal language. Inhetveen (1983: 1) describes this proceeding as a three-step-method.

The achievement test of the present study referred to both these prerequisites of a comparative study as follows: First of all, the ortholinguistic formulated alternatives of the test base on an explicitly introduced terminology. In the item of table 2 that are the term 'exertion of strength', its introduction as a two-place predicator and the formalised definition '$F(K_1; K_2) \leftrightarrow F(K_2; K_1)$' for 'body K_1 exerts strength (F) towards body K_2 and in reverse'. Secondly, in the course developed for the study the learners are introduced to the terminology based on the principles of the three-step-method. At the beginning the learners are free to operate with provided physical devices and to describe their activities with own words such as: 'The second car got a push through the first car'. At the end of a successful course the learners can apply formalisations like the ones of the principal of reciprocal exertion of strength to describe what had happened.

It is recommended to differentiate between implicit and explicit use of the two levels of language. The ortholinguistic terminology is always used explicitly to avoid misunderstandings in empirical studies, the common language explicitly only if required by the objective of a study. Comparing learning outcomes it is recommended to differentiate between three objectives.

- Recommendation no 1: If learning results in different cultures are to be compared with a multiple-choice test, then all alternatives of each item should be formulated exclusively with the help of a ortholinguistic terminology. An example are the alternatives (b) and (c) in table 2.
- Recommendation no 2: To measure the learning outcome the multiple-choice items should have life worldly as well as ortholinguistic alternatives (as the alternatives (a) to (d) in table 2). In pre-test-post-test-experiments with test groups and control groups increasing ortholinguistic formulated solutions describe the learning-growth.
- Recommendation no 3: An analysis of the starting point of learning leads to the categorisation of the life worldly expressions of the learners. Difficulties arising from the ambiguousness of common language texts can be mastered by an ortholinguistic interpretation procedure (compare Krope, 1997).

On Validity

> In selecting or constructing an evaluation instrument the most important question is: To what extent will the results serve the particular uses for which they are intended? This is the essence of validity (Gronlund, 1971: 75)

Three types of validity are commonly used in educational measurement. They are the content validity, the criterion-related validity and the construct validity (compare table 6).

Table 6.
Types of validity, general meaning and methods to obtain validity

Type of Validity	Meaning	Procedure
Content validity	How well the test measures the subject-matter content and behaviors under consideration	Compare test content to the universe of content and behaviors to be measured
Criterion-related validity	How well test performance predicts future performance or estimates current performance on some valued measure other than the test itself	Compare test scores with another measure of performance obtained at a later date (for predication) or with another measure of performance obtained concurrently (for estimating present status)
Construct validity	How test performance can be described psychologically	Experimentally determine what factors influence scores on the test

Table 6 according to Gronlund (1971: 78)

What is the result of the study carried out in Gdańsk and Kiel concerning the validity? An experiment as noted in table 6 has been used as method of determining the construct validity of the mechanics test. It turned out that language is one factor of influence. The improper application of common language and scientific language expressions tends to make test results invalid for their intended use.

References

Bachelard G. (1978), *Die Bildung des wissenschaftlichen Geistes. Beitrag zu einer Psychoanalyse der objektiven Erkenntnis*, Frankfurt.

Bloom B. S. et al. (ed.) (1973), *Taxonomie von Lernzielen im kognitiven Bereich*, 2nd ed., Weinheim/Basel.

Böhme G. (1979), *Die Verwissenschaftlichung der Erfahrung. Wissenschaftsdidaktische Konsequenzen*, [in:] G. Böhme, M. Engelhardt (ed.), *Entfremdete Wissenschaft*. Frankfurt, pp. 114–136.

Domnick J., Krope P. (1972), *Student und Latinum. Untersuchung zum Bestand und Bedarf an Lateinkenntnissen bei Studenten*, Weinheim/Basel.

Gronlund N. E. (1971), *Measurement and Evaluation in Teaching*, 2nd ed., New York–London.

Husserl E. (1976), *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie I/1 (Allgemeine Einführung in die reine Phänomenologie [1912/1913], edited by K. Schuhmann)*, Den Haag.

Inhetveen R. (1983), *Konstruktive Geometrie. Eine formentheoretische Begründung der euklidischen Geometrie*, Mannheim etc.

Janich P. (2001), *Logisch-pragmatische Propädeutik. Ein Grundkurs im philosophischen Reflektieren*, Weilerswist.

Kamlah W., Lorenzen P. (1973), *Logische Propädeutik. Vorschule des vernünftigen Redens*, 2nd ed., Mannheim etc.

Körner St. (1980), *Wissenschaft*, [in:] J. Speck (ed.), *Handbuch wissenschaftstheoretischer Begriffe*, volume 3, Göttingen, pp. 726–737.

Krope P. (1997), *Die Konstruktive Interpretation. Ein Beitrag zur Verbesserung der Objektivität verbaler empirischer Aussagen*, [in:] J. Thonhauser, F. Riffert (ed.), *Evaluation heute – zwölf Antworten auf aktuelle Fragen*, Braunschweig, pp. 187–198.

Krope P., Wolze W. (2003), *Equal and Unequal. Ortholinguistic Basics of the Comparative Method*, Kiel, unpublished.

Lorenzen P. (1987), *Lehrbuch der konstruktiven Wissenschaftstheorie*, Mannheim etc.

Merleau-Ponty M. (1966), *Phänomenologie der Wahrnehmung*, Berlin.

Mittelstraß J. (ed.) (1980), *Enzyklopädie Philosophie und Wissenschaftstheorie*, volume 1, Mannheim etc.

Mittelstraß J. (ed.) (1984), *Enzyklopädie Philosophie und Wissenschaftstheorie*, volume 2, Mannheim etc.

Mittelstraß J. (ed.) (1995), *Enzyklopädie Philosophie und Wissenschaftstheorie*, volume 3, Stuttgart/Weimar.

Mittelstraß J. (ed.) (1996), *Enzyklopädie Philosophie und Wissenschaftstheorie*, volume 4, Stuttgart/Weimar.

Postlethwaite T. N. (1988), *Preface*, [in:] T. N. Postlethwaite (ed.), *The Encyclopedia of Comparative Education and National Systems of Education*, Oxford, pp. XVII–XVII.