PROBLEM SOLVING BEYOND THE LOGIC OF THINGS

Textual and Contextual Effects on Understanding and Solving Word Problems

Kurt Reusser

University of Berne, Switzerland

Address:

Abteilung Paedagogische Psychologie

Waldheimstrasse 6

CH-3012 BERN

Paper to be presented at the annual meetings of the American Educational Research Association, San Francisco, April 1986. The writing of this paper was supported by a fellowship from the Swiss National Science Foundation to the author, and by the Institute of Cognitive Science in Boulder, Colorado.

Abstract

A little recognized topic in the psychological and educational literature of problem solving is the linguistic and extra-linguistic or social-cognitive structure of problem presentation contexts. Word-, story- or text-problems, presented in classroom contexts, represent specific textual and situational patterns of a certain grammaticality. To verbally present a problem to a student in an educational setting always means to somehow organize a fact for the attention of a problem solver. There is the specific structure of the problem text itself by which situations, processes, actions, number relations are implicitly or explicitly expressed, questioned, commented, and there is also the stimulative nature of the pragmatic and social psychological context which shapes the student's textbook-problem solving behavior over a long time.

The present paper outlines and discusses the results of several studies showing that

- subject-matter related to or factual attitudes toward a problem frequently don't play an important part in the problem solving efforts;
- students often solve problems - correctly - without understanding them;
- directionality and the goals of too many problem solving processes are so strongly anticipated by means of various textual and contextual cues that one can hardly speak of the solution any longer as of a genuine achievement of the problem solver;
- false contextual expectations can lead to abstruse errors of understanding and to weird solution attempts.

The experiments also indicate that students can become sensitive and skillful in perceiving and capitalizing on subtle textual and contextual signs pointing to the solution or anticipating its pattern. Moreover, most textbook problems, which are seldomly ill-defined, insolvable or carry irrelevant or other complicating information, and which are almost never presented without an informative question, let students get accustomed to certain courses of processing where a simple fact like whether an equation works out evenly or doesn't, can stop the process or push it further.

Problem Solving 'Beyond the Logic of Things': Textual and Contextual Effects on Understanding and Solving Word Problems

Children learn to deal with text-problems as soon as they enter the highly structured setting of formal schooling in the classroom. It is very likely that over many years the kinds and types of problem texts that are used in school, as well as the pragmatic-situative environment of their solving shape the students concept and style of problem solving. Nevertheless, the structure of commonly used problem texts as well as of problem presentation contexts are barely recognized as topics in the psychological and educational literature of text-, word- or story-problem solving.

Problem solving is widely seen by psychologists and teachers as a process of analyzing a task situation by following its internal factual logic, or as Wertheimer (1945, p. 33) put it, "the inner requirements of the situation". The tradition of "insightful" problem solving first arose with Gestalt psychologists such as Duncker (1935) and Wertheimer (1945), and it has come to life again in cognitive science with the stress on the role of knowledge and understanding in problem solving. While for many decades psychologists discounted the phenomena of insight, planning and understanding, cognitive psychologists nowadays have adopted a quite different attitude towards thinking and problem solving. This attitude comes quite close to the views that Selz (1922), Duncker and Wertheimer had, or had at least anticipated decades ago. Building adequate problem representations, goal-directed planning, inferencing and elaborating by using one's world knowledge, testing hypotheses, applying heuristics and comprehension monitoring are seen as basic operational building blocks of problem solving, as well as of the teaching of thinking skills (cf., e.g., Chipman, Segal & Glaser, 1985; Nickerson, Perkins & Smith, 1985). This paper will not argue against this view of problem solving at all. Instead, the perspective that we are going to illustrate, is intended to complement a view of problem solving that may have been over-idealized.

To present a problem to a student in a classroom or an examination setting means to work up, to pose a factual or fictional situation for the attention of a problem solver. First, there is the wording or specific structure of the problem text itself in which situations, processes, actions, number relations are implicitly or explicitly expressed, questioned, commented upon, (not) excluded and finally, anticipated. Problem texts are grammatical in the many subtle ways they signal paths and goals pointing to the solution pattern and putting the student on the right track. Second, there is the presentational setting, the tacit structure of the pragmatic or situational context that can provide quite significant hints to the solution of a specific task, and that also shapes the student's textbook-problem solving behavior. An important piece of this "context-knowledge" of even very young students concerning textbook math problems is that the problems must always make sense, that they are always solvable, that they work out neatly, that they usually don't contain irrelevant numerical information (i.e., everything that is numerical is relevant), that everything that is relevant is mentioned in the task, and that the explicit problem question - which always accompanies the task - is a reliable guideline in imposing a mathematical perspective on the task or in anticipating (SELZ, 1922) the "operation gestalt" of the solution.

The present paper outlines and discusses the results of several experiments and thinking aloud studies that show how linguistic and extra-linguistic or situational factors facilitate or impede the comprehensibility and solvability of problems. It will be shown how

- strategies following 'the logic of things', or factual attitudes towards a problem frequently don't play an important part in the problem solving efforts;
- students often solve problems - correctly - without understanding

them;

- directionality and the goals of too many problem solving processes are so strongly anticipated by means of various textual and contextual cues that one can hardly speak of the solution any longer as a genuine achievement of the problem solver;
- ill-defined or false contextual expectations can lead to abstruse errors of understanding and to weird solution attempts.

Irrelevance and Pre-knowledge in the Missile Task

The phenomena I am going to discuss first stems from an observation in an experiment which was originally aimed at understanding the effects of irrelevant numerical and episodic information in a text problem. The main question was how irrelevant information would affect problem difficulty depending on what kind of instructional aids were provided. We will not report the whole experiment here, but concentrate on an interaction between the formulation of an applied physics problem and two levels of pre-knowledge.

In the first version of the experiment, four versions of the missile task (Table 1), manipulating the extent and the quality of irrelevant

---

Insert Table 1 about here

---

information, were given to groups of 13-14 year old high school and to 18-19 year old college students. As we hypothesized from previous observations irrelevant scriptal information was easier to recognize as irrelevant and to eliminate than was irrelevant quantitative or numerical information (Figure 1).

Insert Figure 1 about here

A rather astonishing fact, however, is also revealed in Figure 1: Considering the frequency of correct solutions, the college students with some background in math and physics weren't any better than the eighth and ninth graders in version T4, the task with the most irrelevant numerical information. This observation led us to look closely at the student's written solution protocols. What we found there led us to conduct a second experiment, this time only comparing T1 and T4, now predicting that the college students (n=39) and the high school students (n=57) should differ in T1 but not in T4. This is exactly what happened. The result of this comparison is shown in Figure 2.

Insert Figure 2 about here

First of all, the almost opposite tendency with respect to T4 is due to an effect specific to one group of college students who participated in the experiment shortly before graduation: While 11 students solved T1 and 1 student didn't, only 1 student solved T4 and 10 didn't. It became clear what had happened when we looked at the solution protocols and the retrospective reports of some of the college students. What most of the college students did, activated by the giving of the distance and the launch-time delay and their more elaborated physics knowledge, was to construct a far more complex problem space, not only more complex than was required by the task, but also far more complex than their mathematical abilities could handle. The college students tried hard to do something impossible, namely to get a grasp of the mathematica description of the trajectory of the missile. They typically interpreted the proposition [APPROACH(A,B,DIRECTLY)] as implying a curved trajectory, while the high school students took it for granted that it was a straight line. This lead to different notions of distance. A minority of college students concluded the task was unsolvable or ill-defined. Comments and questions concerning launching angles and shape of the trajectory, as well as the drawings and the many complicated calculations found in the protocols, clearly indicate the complexity of the problem spaces that the college students - so to speak the expert subjects in the experiment - were forming. The college students knew (just as well as the high school students knew the opposite) that they should be able to come up with the calculation of a trajectory. However, the task is simply not specific enough to permit these calculations. Since, as we will show later, it doesn't seem to be at all easy to recognize a given task as being unsolvable, the majority of the college students tried hard to understand mathematically what happens in the task, with only a few finally seeing how the task could be solved under very simple assumptions.

The high school students were in a more comfortable situation than their more expert colleagues. Because of their modest physics and mathematical knowledge, they could not form a complex view of the physical task space or carry out demanding mathematical calculations. Their minimal knowledge didn't lead them to perceive the task as as complex as the consideration of the information, intended to be stripped off as irrelevant, would require. Those students who thought of of a curved trajectory, of the curvature of the earth or of the value of some launching angle, soon dropped these elements from his representation, because he knew that he had never calculated them before, and that he would never be able to do so. The high school student - as well as the college student on a more expert level - was not capable of seeing more structure in the given task as his assimilative knowledge base was able to interpret: Distance ... time ... velocity ... there may only be a single formula available, v = d : t , and some knowledge how to transform it algebraically ... -> d = vt ... -> t = s:v ... and nothing more is needed than to try a little, to look for an instantiation of the formula that works out and that fits the context of previous experiences and demands. Put another way: The lack of physics and mathematical knowledge led the students eliminate things which were of no use and for which no calculational methods were known. The smart high school student, who may even know how the problem situation could be understood in a more complex way - with a good feeling for the demands for text problems - refrained from maximizing his semantic in-depth processing. He had probably often experienced that in most text problems in the classroom context there is a sort of "prestabilized harmony", a certain match between subject-matter complexity and the mathematical knowledge required for solving a problem. From the didactic point of view of those who design mathematical text problems, there is often a hidden conflict between, on the one hand, the real world complexity of task situations and mathematical instruments required to cope with them, and, on the other hand, the modesty of the student's actual domain-specific and mathematical knowledge. There is no easy way out of this dilemma, and it is also hard to see what more general impact this difficulty has on student's problem solving behavior outside the classroom, where no designer of a situation problem, a priori, guarantees solvability or an easy match between one's own (mathematical) resources and the semantic requirements of a task.

Text and Context Related Difficulty Misjudgment in the Bicycle Task

The problem text is just one aspect of the input in problem solving, as well as being just one of the guiding forces. Whoever observes students in classroom and homework situations can find again and again how few common textbook problems force students to do an in-depth semantic analysis, how many students are striving to orient their problem solving in response to a multitude of indicators in the context of the factual problem structure - and how they succeed. The experiment that we are going to report isolates a factor "estimated or assumed task difficulty". This was done on the basis of observations indicating that students, often before or in place of a thorough analysis of the real content of a task, use context information to assess the direction and difficulty of a solution. The single task we employed in this experiment allowed us to manipulate the factor "estimated task difficulty" in two ways: as a text factor and as a context factor. Four different settings of the cyclist task were presented to ninth graders and to college students.

Task description:

Table 2a shows the simple (S) and the more complicated (C) text version of the cyclist task. Both problem texts were accompanied by the sketch in Figure

---

Insert Table 2a and Figure 3 about here

---

3 and the basic distance-time formula. The formula was provided so that the high school students would not fail merely because they could not reproduce the formula from memory. Distance-time tasks like the one we employed belong to the basic math curriculum of the junior high school. The task is seductive because it often puts people off the scent by letting them simply average the three partial speeds. This temptation, which we found by presenting the task to a few subjects, seems to be smaller or bigger according to how the velocities in the task are expressed. The two textual versions take these observations into account. While S, a suspiciously simple looking version, might warn the average problem solver to be cautious with regard to a (v1+v2+v3)/3 solution ("What's the difficulty here?"), the reassuringly complex version C induces the subject to expect rather the opposite: The problem solver finds at least one 'difficulty' in the calculation of the partial speeds v2 and v3, and since, moreover, the following calculation (v1+v2+v3)/3 works out evenly, he may think the problem is solved. Further, two co-textual additions were constructed with the function of explicitly inducing an expectation of either high (IDH) or low (IDL) task difficulty

(Table 2b). The difficulty induction

Insert Table 2b about here

was made by associating the task to school-specific types of examinations. Different additions had to be developed for both the high school and the college subjects. However, the relation between IDH and IDL is assumed to be the same for both groups.

Four task settings resulted from the combination of the text versions with the co-text additions:

S(IDL) : The induced low difficulty and the simple text version support each other. Quite a large number of (v1+v2+v3)/3 -solutions are to be expected.

C(IDL) : In contrast to S(IDL), v2 and v3 cannot be directly read off from the problem text but have to be worked out before the main calculation.

This 'element of difficulty' - in interaction with the IDL-addition - will probably make many subjects rather careless. There are even more errors to be expected than with the first setting - presumably the most of all settings.

S(IDH) : This version can be viewed as the complement to C(IDL).

In the face of the high difficulty context, only a few subjects ought to accept a (v1+v2+v3)/3 solution. Therefore, in this condition, the most correct solutions are expected.

C(IDH) : Since for many high school students, the preliminary calculations of v2 and v3 are expected to be of significant difficulty, it can be assumed that, for many of them, this difficulty is taken into account sufficiently by the IDH context. Others - and presumably most of the college students - will look for an additional difficulty and do some deeper semantic processing. Altogether we still expect a fairly large number of erroneous solutions, but significantly fewer than under the conditions C(IDL) and S(IDH).

Hypotheses:

Three main predictions underlie the experiment.

H 1 : For both text versions, a higher frequency of correct solutions is expected for the condition of an induced high-difficulty context than for the low-difficulty context.

[f+] S(IDH) + C(IDH) > [f+] S(IDL) + C(IDL) => [IDH > IDL]

H 2 : In case of the induction of an equal difficulty, a higher frequency of correct solutions is expected under the S- than under the C-Condition.

[f+] S(IDH + IDL) > C(IDH + IDL) => [S > C]

H 3 : The highest frequency of correct solutions is expected for the condition S(IDH), the lowest for C(IDL).

[f+] S(IDH) > C(IDL) ;

[f+] S(IDH)max ; [f+] C(IDL)min

Results:

68 Bernese junior high school students (median age 15) and 51 college students (median age 19) participated in the experiment which was run in groups, by two experimenters during normal class hours. The written solution protocols - Figure 4 shows some examples - were first assigned to two restrictively defined categories:

* <u>Correct solutions</u>: Subjects were assigned to this category if they first calculated the times for each of the three distance-segments, summed up the times and then put the sum as the denominator into the basic formula (cf. example A in Figure 4).

$$v(\text{average}) = \frac{3s}{\underbrace{\frac{s}{v1}}_{t1} + \underbrace{\frac{s}{v2}}_{t2} + \underbrace{\frac{s}{v3}}_{t3}} = \frac{30}{\frac{10}{9} + \frac{10}{48} + \frac{10}{18}} = 16 \text{ km/hour}$$

* <u>Incorrect solutions</u> corresponding to the pattern "average the partial speeds" (aps-pattern; cf. example B in Figure 4).

$$v(\text{average}) = \frac{v1 + v2 + v3}{3} = \frac{9 + 48 + 18}{3} = 25 \text{ km/hour}$$

This analysis turned out to be too rigorous, especially for the high school subjects, where only 10% solved the problem completely correctly. Moreover, 30% of the protocols couldn't be classified because of all sorts of errors. Therefore, three raters analyzed the protocols again along the following lines: Whoever's solution attempt clearly showed the calculation of the partial times ti - even if this was done by the wrong formula (v/s instead of s/v ; cf. examples C in Figure 4) - was assigned to the correct

---

Insert Figure 4 about here

---

solution category. Due to the very specific structure of the task (one salient erroneous solution), virtually all subjects could be classified either as solvers, according to the above (weaker) criterion, or as non-solvers of the average-partial-speed kind. The frequency data for the correct (+) and the incorrect aps-solution (-) attempts of both subject groups are shown in the figures 5a, 5b, and 5c. We will now discuss each of the three hypotheses.

Insert Figures 5a, 5b, 5c about here

---

* Effects of difficulty induction [H 1 : f(IDH) > f(IDL)] : The reader will have noticed that the high school students produced rather few correct solutions (floor effect) whereas the college students showed only a few aps-solutions (ceiling effect), which weakens the statistical analysis. But this was the price we had to pay for using the same task in on two very different school levels. The frequency data confirm our first hypothesis that there are more correct solution attempts under the IDH- than under the IDL-condition (Figure 6). The difference becomes significant for the

---

Insert Figure 6 about here

---

college students if we add some subjects. There were no differences in solution times between IDH and IDL.

* Effects of problem formulation [H 2 : f(S) > f(C)]: The direction of the S/C-comparison for both groups is consistent with the previous prediction, but only the combination of the groups shows a statistically reliable difference (Figure 7). When we later tested 50 more college students comparing only S(IDL)

Insert Figure 7 about here

with C(IDL), we found a reliable difference in the expected direction ($p<.025$). So, even an expert-like group, for which simple algebraic transformations were by no means a source of difficulty, was sensitive to the induction of a low-difficulty context. What was highly affected by the problem formulation were the solution times of the high school students, but not of the college students - which makes sense: Even if the algebraic transformations in C were not an obstacle for solving the problem, it took the high school students a lot of time to do them, while the more expert students - due to their much better subroutine- skills - showed no difference. Further, both groups also needed more time to come up with the correct solution than with the incorrect one.

* Maximum- and minimum-frequency [H 3 : S(IDH)max ; C(IDL)min] : The presentation of the very simple task associated with a high-difficulty context showed the most correct solutions, the association of the enriched task with

Insert Figure 8 about here

an induced low-difficulty context the fewest. H 3 can be confirmed for both groups (Figure 8).

Discussion:

Three out of four high school students and one out of five college students were wrong in their solution of the cyclist task, succumbing, as it were, to its seductive textual and co-textual features. The results support our view that a contextual orientation ought to be seen as a constitutive factor of comprehension in word problem solving. Students of very different levels of expertise interact with a problem text not only in a fact- or subject-matter- related way, but also orient themselves - as in the present experiment - towards a co-textual context. The most striking difference in the data is that between the conditions S(IDH) and C(IDL). Is this difference due to a direct effect of the differences in the formulation of the task, or is it to be explained as an indirect effect due to an interaction of the textual characteristics with the difficulty-induction characteristics? We can clarify this issue by analyzing the C-version solution protocols with respect to the question how well the conversions of v2 and v3 were done under the conditions IDH and IDL. What Table 3 shows is that a simple and direct impact of the conversions of v2 and v3 on

Insert Table 3 about here

the solution process - under both IDH and IDL - can be excluded: Neither are there ID-related differences in correct conversions within both subject groups, nor is there a significant difference between high school and college students. The observed difference between the groups would be even smaller if one didn't take into account the mere calculation mistakes of the high school students in the course of working out v3. Because there is a difference between C(IDH) and C(IDL), in any case, we have to conclude that there is an indirect effect of a (per se) very simple conversion operation required by the C-version of the task. We can think of this effect as follows: The induced difficulty cause the students to expect complications. A high school student under IDL hardly expects any serious difficulty. Nevertheless, the conversion of v2 and v3 provides at least a minor difficulty, but there is no reason to seek further complications. There is also the (disarming) fact that the incorrect solution works out evenly, and therefore most high school students -

and also 31% of the college students - become careless and stop without any further epistemic control. - The situation is much different when a IDH context is given: The student expects a rather complicated situation. He is probably more observant, and EINGESTELLT (Luchins, 1942) on a task with a certain amount of complications. At least some of these expectations are fulfilled - probably at little bit more for the high school than for the college students: The conversions have to be done without any mistakes, which takes at least the high school students some time, but doesn't provide a significant difficulty for them. So, there remains an unfulfilled expectation; there still must be another difficulty: the student carefully recalls (vergegenwaertigt) the problem situation again, and - in some cases - succeeds in finding the crucial point. - The student may be even more unsatisfied under the condition S(IDH). The task elicits high expectations of complications which may never be fulfilled. The expectations are not met by any conversions or the required preparatory operations. The contradiction between the activated level of expected task difficulty and the initial representation of the problem or the first problem model is maximum, and therefore challenges the student to more deeply process the problem, with the result that under S(IDH) the most correct solutions are found. So, the additional calculations in the C-versions impede the solution process not by imposing any substantial difficulty on the task, but by inducing, as a moderating variable, in-depth processing.

Therefore, the problem solver proves himself to be not only one who is driven by "the desire ... to go on from an unclear, inadequate relation to a clear, transparent, direct confrontation - straight from the heart of the thinker to the heart of his object, of his problem" (Wertheimer, 1945, p. 236), but also one who, after having built an initial problem model by using all available textual and contextual clues, allocates the amount of resources in processing time and energy he expects to use. One could put it this way: A problem solver, after or while reading a specific task, allocates resources to be used in a solution attempt, or opens, so to speak, a sort of "cognitive processing energy account", against which all steps of processing get recorded. If there is nothing to "debit" the account - given a high "energy credit" - as under S(IDH), the problem solver will seek to work himself deeper into the material as if, as under C(IDL), his expectations were more or less fulfilled. In this view, to gain a deeper level of comprehension also means to activate or to strengthen the control functions. There were quite a few subjects under IDH who first worked out (v1+v2+v3)/3 , but later discarded this solution.

The Classroom Context of Problem Solving

How a problem is understood and solved, and how difficult it is, depends in the first place on its wording as a task. However, as the following examples will show, even the situational context within which a problem solving process takes place may have a significant influence on the understanding and solving of a problem.

How the Solving of a Problem Can Become a Joke

One can study the context- or situation-dependence of problem solving processes by looking at what impact the negation of the situation has. Suppose a college student gets the following task at an oral examination in physics:

> "Show how one can measure the height of a skyscraper with the help of a barometer."

Suppose further that the student answers correctly:

> "One can determine the height of the building by reading off from the the barometer the air pressure difference between road and roof. Air pressure decreases by 1 Torr (= 1 mm Hg) approximately every 30 feet."

It could easily happen that the candidate doesn't know the answer. This would probably result in a bad grade.

Sensitized by my interest in context phenomena in problem solving, I came across a short text, reporting how a candidate answered this question in a completely different way.2) He produced not only one but a whole series of answers - not to the pleasure of the examiner - according to the report.

(a) "You take the barometer with you to the top of the roof, tie it to a long rope and lower it to the road. Then you pull it back up and measures the length of the rope. This length corresponds to the height of the skyscraper."

(b) "... or you take the barometer outside on a sunny day, put it on the ground and measure its height and the length of the shadow. Then you determine the shadow of the skyscraper and calculate the height of the building with a simple proportional equation."

(c) "You take the barometer with you going up the stairs of the building. In the course of this you mark the wall in 'barometer-units'. The only thing you have to do afterwards in order to get the height of the building, is to count the 'barometer-units'. This is, of course, a very clear but rather crude method."

(d) "You take the barometer to the top of the building. Then you lean out over the edge of the roof. You drop the barometer and measure the falling time with a stop watch. Then you determine the height of the building by the law of falling bodies: d = 1/2 gt2 ."

(e) "If you were interested in a more subtle method, tie the barometer to a rope and let it swing as a pendulum. You determine the value of g (gravitational force in the formula T = 2 |1/g) on street and on roof level. Then you can work out the height of the building from the difference between g1 and g2."

(f) "Finally, if you don't want me to commit to a physics solution, then there still are many more possibilities. For example, you could take the barometer and knock on the janitor's door. If he answers the door, the you speak as follows: 'Dear janitor, I have here an exciting barometer. If you tell me the height of the building, then it's yours."

What remains to add is that, of course, the candidate also knew the 'correct' solution. - What happened? Obviously, a problem solving process suddenly became a sort of a funny joke. I hope that everyone will feel sympathy for the refreshing originality of the candidate. The little story sheds light on the examination context as a familiar problem solving setting, and it illuminates very nicely what can happen if its more general and task-specific constraints are disregarded by the candidate. One can look at what happened from the point of view of '<u>functional fixedness</u>', a phenomenon that was discussed first by Duncker (1935) and many others after him who studied the effects of "Einstellung" (Luchins, 1942; cf. Greeno & Simon, 1984). From the functional-fixedness point of view, quite a few of the candidate's solutions would have to be seen as very <u>difficult to retrieve</u>, because they abstract, with regard to the functional character of the object barometer, rather <u>remote and unusual features</u>. A barometer is an instrument for measuring air pressure. This is its <u>primary function value</u>. It would probably be very hard to come up with application-contexts where the barometer is used as a pendulum, a shadow-producing object, or a bribe for caretakers. Whoever is able to see a barometer under such a variety of only faintly moulding features, demonstrates creative behavior, though it presupposes a context in which the behavior is also perceived as original or creative. The typical examination context is not such a context, and so there was much argument about the value of the candidate's proposed solutions in our little story. And one must almost certainly assume that many examiners, were they exposed to similar situations, would feel insulted, provoked, or made a fool of. But why, really? I am going to try first a more general answer and then explore two guesses, which follow from it.

Oral examination situations are <u>behavior settings</u> (Barker, 1968) with a defined structure. The external course of events is mostly fixed, similar to a script (SCHANK, 1977). The actors in the situation play roles with well-defined expectations. They are, as it were, partners who have entered a - temporary limited - <u>speech act contract</u>. "For interaction to succeed, (both) participants must agree in their social situation definition" (Leodolter & Leodolter, 1976; cited from Forgas, 1985, p. 19). This "social contract" (cf. Mead, 1934) includes, for the candidate and the examiner, specific speech act obligations and rights. Greatly simplified:

- The examiner is entitled and obliged
  . to question the candidate on a previously defined topic and to present problems to be worked out by the candidate,
  . to judge the candidate's responses using criteria related to the subject-matter.

- The candidate is entitled and obliged
  . to prepare himself / herself to be examined about topics previously defined or agreed upon,
  . to obediently take the posed questions and to answer them after a short period of thinking.

It is not my point to elaborate this oral examination context, even if this were possible. It is only important here to see that such a context exists, and that we can assume that the candidate, the examiner and the reader of the story know it very well. It is this examination context, whose inherent obligations are systematically and intelligently ignored by the candidate above. Not that the candidate wouldn't live up to the expectations of the external examination script. It can on the contrary be assumed that the candidate

- was very polite,
- gave the examiner, while articulating his responses, a well-meaning and zealous impression,
- didn't miss any questions.

Basically, there is only a single behavior expectation that gets systematically and consciously negated by the candidate: The expectation to understand and to answer the question in a situationally defined way, which means that a solution is only acceptable if

- it can be regarded as, in a certain sense, intellectually demanding,
- it is founded on knowledge of the physics topics previously agreed upon,
- the barometer is regarded in its central function value as an instrument for measuring air pressure.

Because of the lack of explicitness of these expectations, the candidate does not neglect the letter but the spirit of the examination context. He irritates the examiner the same way that he amuses the reader of the story. With his solutions, none of which is based on the barometer's central function, but on remote functional and dispositional features (extension, weight, exchange value) of the object, the candidate reduces the examination to absurdity.

Stimulated by this little story, I got interested in two questions:

First, would other readers also find the story funny the same way I did? Could one possibly find a consistent rank order of solution-funniness as a consistent pattern of funniness judgments?

Second, could we predict the funniness of the solutions by the degree of their deviation from the pattern of expectations implied by the oral examination context and anticipated by the examiner's explicit question?

We predicted the following task-specific rank order of funniness (listed in reverse order). (1) means that this solution is considered to be the funniest:

(7) The correct solution: It corresponds entirely with both the textual and the contextual anticipations. Nobody will perceive this solution as funny. The function value of the solution is identical with the function value of the barometer.

(6/5) The pendulum (e) and the shadow (b) solution: Both are demanding from a math or physics standpoint, even if they do not employ the barometer in its essential function at all but in some of its merely accidental properties. Both solutions require thinking and some serious knowledge. They are neither dubious nor just convenient.

(4) The barometer as a unit of measurement (c): Admittedly, the solution is not very elegant, rather hard-working, but nevertheless simple and efficient. Besides that, the candidate negates or fails to meet the expectation of presenting a physically demanding solution.

(3) The rope-solution (a): The solution is not based on domain-specific knowledge or on skill or on staying power. It is convenient, yet it can be worked out in a lying position, and it undoubtedly leads to a correct and precise result.

(2) The janitor-solution (f): It simply defies any description of an intellectually honest solution in an examination context. The candidate not seeking the solution himself, but rather tries to buy it for the price of the "solution instrument": The epitome of a bone-idle, dishonest solution.

(1) The free-fall solution (d): This solution requires some knowledge of physics, which speaks against considering it the funniest. Unlike the the case of the 'janitor-solution', where one might think that an unintelligent or desperate problem solver could see no other way out but to sell the solution instrument for the solution, no argument of mere convenience or stupidity can be brought to bear in this solution. The candidate, who knows the law of falling bodies, who proposes to

work the solution not by letting fall, e.g., a stone, but by the irreversible destruction of the solution instrument, acts negligently in a way that transcends even the specific examination context - by the deliberate destruction of something of value.

We presented the seven solutions to 8 university students and asked them to put them in a funniness rank order. Table 4 shows two things: a high

Insert Table 4 about here

correspondence between empirical and predicted rank order, and a high consistency within the empirical rank order (coefficient of concurrency after Kendall, $W = .70$; $p<.001$).

Not infrequently, exam questions require, like many text problems, a certain sensitivity or cleverness in 'reading off' the intentions, anticipations and expectations from the text and the context of a problem. Whoever has this sensitivity for context, together with intelligence and knowledge, and who deliberately doesn't take it into account, can subvert a problem situation until it develops into the purely comical.

The Authority of Contexts and its Impact on Comprehensibility

The Swiss writer Peter Bichsel in his third Frankfurt Poetics lecture (1982, 49) remarked about his reading of Goethe's "Joseph":

> Maybe I would have broken off my reading if the author had been unknown to me. For example, I could have stopped reading because I could have assumed that the book was sort of sanctimonious, or, if you want, simply a book for people and not a book for literature. You may well interpret that as snobism. But my literary judgement is dependent of the context: I read that book in the context of Goethe and in the context of German literature.

Whoever is confronted as a student in a classroom, expects something comprehensible and solvable. And he knows that he is expected to produce an answer - even when there may be none. From the standpoint of divergent fantasy, one might call the sense-seeking behavior of students creative vis-a-vis nonsensical material or unsolvable problems, and sometimes there really are some such solutions. But this isn't the rule. Most of the time one encounters rather questionable or even ugly ways in which students try to understand and solve a problem a tout prix, because of the characteristic/moulding/shaping factor of context and its dubious impact. I would like to call that questionable impact the lack of intrinsic cognitive processing. Five examples:

Example 1

> I asked the class: 'Are you sure that this result is really correct?' Most of the pupils were plainly dumfounded by the question, surprised that it should be asked. Their attitude was clearly: 'How can you expect us to question the solution you have given us?' The question was strange to them, it touched the very essentials of what school, teaching, learning meant to them. No answer. The class was silent (Wertheimer, 1945,26).

Before Wertheimer asked the students this, he showed them how one can work out the area of a parallelogram with a very troublesome, unpleasant and senseless method - but leading to a correct result. Wertheimer comments further:

> Let the reader consider whether he has not often learned things in school that way. Isn't it the way in which perhaps you have learned differential and integral calculus? Even theorems of plane and solid geometry? Of course you had good reason to feel that the teacher was teaching sensible, serious things you had to learn. But did you have the possibility of another kind of learning, of really grasping? Could you do anything but put up with and submit to the teacher's demonstration, step by step, when you were unable to see why he did just this, then that? Could you help just following obediently as the steps dropped out of the blue? (26)

While following an explanation or a demonstration, students often don't become really challenged enough to understand what is presented to them, and to evaluate it by means of their own criteria of consistency and comprehension quality. Such criteria might not yet be available, but what does teaching do to help these develop?

The following examples were collected over the last few years. They may