STUDENTS' OVERRELIANCE ON LINEARITY: AN EFFECT OF SCHOOL-LIKE WORD PROBLEMS?

Wim Van Dooren [1][2], <u>Dirk De Bock</u> [2][3], Dirk Janssens [2] and Lieven Verschaffel [2]

[1] Research assistant of the Fund for Scientific Research (F.W.O.) – Flanders
[2] Center for Instructional Psychology and Technology, University of Leuven, Belgium
[3] European Institute of Higher Education Brussels (EHSAL), Belgium

Previous research showed students' tendency to improperly apply the linear model when solving non-linear problems about the relation between lengths, area and volume of enlarged figures. Most of these studies, however, were conducted with collective tests containing traditional, "school-like" word problems. The current study shows that students' problem-solving behavior strongly improves when the non-linear problem is embedded in a meaningful, authentic performance task. It is also found that this experience does not affect students' performance on a posttest, where the non-linear problem is offered again as a word problem.

INTRODUCTION

Because of its wide applicability for understanding mathematical, scientific and everyday life problems, linearity (or proportionality) is a key concept throughout primary and secondary mathematics education. Inherent to the attention it receives, however, is the risk to develop an overreliance on the concept: "Linearity is such a suggestive property of relations that one readily yields to the seduction to deal with each numerical relation as if it were linear" (Freudenthal, 1983, p. 267). The tendency to overgeneralise the linear model is repeatedly mentioned in the mathematics education literature, and in recent years it has also been in the focus of systematic empirical research. For example, the phenomenon has been studied in elementary arithmetic (Van Dooren, De Bock, Hessels, Janssens, & Verschaffel, 2005), algebra and calculus (e.g., Esteley, Villareal, & Alagia, 2004) and probability (Van Dooren, De Bock, Depaepe, Janssens, & Verschaffel, 2003).

The best-known (and extensively studied) case is situated in geometry: many students of different ages believe in a linear relation between the lengths, areas and volumes of similarly enlarged geometrical figures, thinking that if a figure is enlarged k times, the area and volume of that figure are enlarged k times as well (De Bock, Verschaffel, & Janssens, 1998, 2002b; De Bock, Van Dooren, Janssens, & Verschaffel, 2002a; Freudenthal, 1983; Modestou, Gagatsis, & Pitta-Pantazi, 2004). A series of studies has shown that even with considerable support (such as providing drawings, instructing to make drawings, or giving metacognitive hints), the large majority of 12- to 16-year old students failed to solve these problems due to an alarmingly strong tendency to apply linearity (De Bock et al., 1998, 2002b; Modestou et al., 2004). Further research showed that the tendency was due to a set of closely

related factors like the intuitiveness of the linear model, shortcomings in students' geometrical knowledge, inadaptive attitudes and beliefs towards mathematical (word) problem solving and a poor use of heuristics (De Bock et al., 2002a).

These last explanatory factors (namely: attitudes and beliefs towards word problem solving and a poor use of heuristics) led us to conduct the study presented in this paper. In most of the previous research the overreliance on linearity was observed in a classical scholastic context by means of collective tests with word problems, i.e. short written descriptions of a problem situation with the task to do some calculations and to write down a short numerical answer. It can be argued that the use of word problems may trigger in students a set of implicit rules and expectations established by the socio-mathematical norms of the classroom setting (Cobb, Yackel, & McClain, 2000; Verschaffel, Greer, & De Corte, 2000). Possibly, the students in our previous studies may not have invested sufficient mental effort in the solution of the problems – assuming that they were dealing with routine word problems –, or may have excluded a number of considerations and problem solving strategies (e.g., checking the viability of a solution by making a sketch of the situation) – assuming that they were not desirable, acceptable or valid in that context. Research evidence shows that students are more inclined to leave their routine word problem solving behavior and include real-world knowledge when the problems are disentangled from their scholastic chains and embedded in more meaningful, authentic "performance tasks" (e.g., DeFranco & Curcio, 1997; Nunes, Schliemann, & Carraher, 1993; Reusser & Stebler, 1997). The current study aimed at investigating whether this would also be effective to break students' overreliance on linearity: Can this tendency be weakened or even eliminated by embedding non-linear problems in meaningful, authentic performance tasks instead of traditional, school-like word problems?

METHODOLOGY

The study was conducted in three steps. First, participants were selected using a pretest. Next, students who made a linear error on the pretest were involved in an individual interview. And third, a posttest was taken of all the interviewed participants. Each part is explained in more detail below.

Selection of participants by pretest

The first step was to select students who were prone to the error under consideration. 93 sixth graders (i.e., five whole class groups in two different schools) solved a pretest that contained six word problems. Five of the word problems acted as buffer items (they were included to avoid revealing the focus of our study). One word problem in the test aimed at detecting whether students tended to give a linear answer to problems about the effect of an enlargement on the area of a square:

> John needs 15 minutes to paint a square ceiling with a side of 3 meters. How much time will he approximately need to paint a square ceiling with a side of 6 meters?

Altogether, 72 students gave a linear answer to this problem (e.g., "3 m × 2 = 6 m → 15 min × 2 = 30 min"). They were involved in the rest of the study.

Interview procedure

Two days after the pretest, these 72 students were taken individually out of the classroom for a semi-structured in-depth interview. During that interview, the students again were asked to solve a non-linear problem, this time about the effect of *tripling* the lengths of the sides of a square on the area of that square. This problem was offered in one of three different ways, depending on the interview condition that the student was assigned to. Assigning students to interview conditions happened by means of matching on the basis of school mathematics performances.

Students in the *S-condition* ("Scholastic" condition, $n = 24$) received a sheet with the following traditional, scholastic word problem:

> Recently, I made a dollhouse for my sister. One of the rooms had a square floor with sides of 12 cm. I needed 4 square tiles to cover it. Another floor of the dollhouse was also a square, but with sides of 36 cm. How many of those square tiles did I need to cover it?

The problem in the *D-condition* ("Drawing"-condition, $n = 24$) was the same as in the S-condition, but this time the sheet also contained a drawing of the small and large figure, as shown in Figure 1.

In the *P-condition* ("Performance task"-condition, $n = 24$), the problem was presented as a "performance task": Students were involved in an authentic problem context with real materials (the small dollhouse floor, 4 tiles and a large dollhouse floor) and were asked to perform an authentic action. The interviewer presented the task as follows:
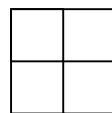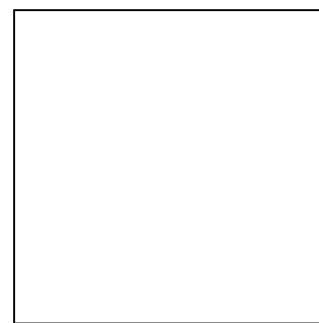
Figure 1: Drawing offered with the word problem in the D-condition

> I have a little sister, and currently I am making a dollhouse for her. Here, you can see the floor of one of the rooms. Can you tell me its shape? [The student tells that it is a square.] Let's measure it. [Student observes that the sides are 12 cm long.] I have some tiles that we can use to cover that floor. Can you do that for me? [Student puts 4 tiles on the small floor.] Indeed, we need 4 tiles to cover this floor.

> I also brought another floor of the dollhouse. As you see, it is also square. Let's measure it. [Student observes that the sides are 36 cm.] In a few moments, we will put tiles on this large floor as well. Now, think about how many tiles you will need to do that, and if you have decided, you can go and get exactly enough tiles from the table over there.

The number of tiles brought was registered as the students' final answer. At the end of the interview, students in the P-condition were allowed to put the tiles effectively on the large floor (and could get more tiles if necessary).

Strictly spoken, for our research goal the design only would need to include a S-condition and a P-condition, but nevertheless the D-condition was included, because students in the P-condition not only received the non-linear problem as an authentic performance task instead of a school-like word problem; the problem presentation in the P-condition involved also visual support, while this support was not present in the S-condition. Including a D-condition – which provided the same visual support as the P-condition, but the same scholastic presentation as the S-condition – allowed us to control for this visual support factor.

All interviews were registered on videotape, and students were asked to think aloud while solving the problem. They were told that they could solve the problem in whatever way they wanted and use all materials available (pen, paper, ruler, pocket calculator and in the P-condition also the small and large floors and the 4 available tiles). When necessary, the interviewer asked some additional probing questions to clarify students' thinking. At the end of the interview, the students were asked to indicate on a five-point scale how certain they were about the correctness of their answer ('certainly wrong', 'probably wrong', 'no idea', 'probably correct' and 'certainly correct'), and to justify this[1].

Posttest

One or two days after their interview, students solved a posttest. Besides five buffer items, it again included a non-linear problem that referred to the same mathematical situation as the pretest item (the effect of doubling the sides of a square on its area):

> Carl needs 8 hours to manure a square piece of land with a side of 200 meters. How much time will he approximately need to manure a square piece of land with a side of 400 meters?

Because this problem situation only slightly differed from the one that students had in the interviews (where the sides of the square were tripled), it could be determined whether the experiences during the interview also had a learning effect. For example, manipulating the materials during the interview in the P-condition could be helpful for the student to solve the non-linear problem on the posttest correctly too.

RESULTS

Individual interviews

Table 1 provides a summary of the answers and the solution time (i.e. the time needed to find an answer after the problem was introduced) of the students in the three interview conditions. The table shows that there was a strong impact of the interview condition on students' answers (Fisher's exact test $p < .00015$):

---

[1] Evidently, P-condition students had to answer the probing questions and the certainty question before they were allowed to put the tiles on the large floor to check their answer.

Nearly all students in the *S-condition* (i.e. 21 of 24 students) erroneously applied linearity to solve the word problem. This confirms – once again – students' very strong tendency to stick to the linear model when solving problems about the area of enlarged figures, as observed in previous studies (De Bock et al., 1998, 2002a, 2002b). Two students committed another error, and only one student found the correct solution (in contrast with the pretest, this student now made a drawing of the problem situation which led him to the correct solution).

In the *D-condition*, the performance was considerably better. Here, 16 students found the correct answer during the interview. In fact, we had expected that the visual support as such would not be helpful for most students, since in previous studies with collective tests (see, e.g., De Bock et al., 1998, 2002b), the provision of ready-made drawings hardly had any effect on students' performance, mainly because students simply neglected them. In the current study, however, many students did actually use the drawing – possibly an effect of being involved in an individual interview context where they felt more obliged to do so – and their solution process clearly benefited from it. Nevertheless, the provision of a drawing was not sufficient to eliminate all linear reasoning: 8 out of 24 students still gave a linear answer to the word problem.

Although presenting a drawing was beneficial for many students, presenting the non-linear problem as an authentic performance task had an even stronger impact on students' reasoning (Fisher's exact test yielded $p = .0412$ for a separate comparison of the D- and P-condition results). In the P-condition, 20 students gave the correct answer, and only 2 students reasoned linearly (and 2 students made another error).

In sum, almost all students in the S-condition made the linear error, which is not surprising considering that they did the same on the non-linear word problem on the pretest. Providing drawings had a positive effect on students' performance, but still one third of the D-condition students made a linear error. Offering the problem as a meaningful, authentic performance task was even more beneficial, since in the P-condition linear errors were nearly absent.

Interviews				Answer on posttest		
Condition	Answer	Freq	Solution time (seconds)	Correct	Linear	Other error
S-condition	Correct	1	120	1	0	0
	Linear	21	50	1	20	0
(*n* = 24)	Other error	2	115	0	2	0
D-condition	Correct	16	139	0	16	0
	Linear	8	61	1	6	1
(*n* = 24)	Other error	0	/	/	/	/
P-condition	Correct	20	76	2	17	1
	Linear	2	29	1	1	0
(*n* = 24)	Other error	2	131	0	0	2

Table 1: Overview of answers and average solution times in each interview condition and of answers on the posttest

A closer look at students' solution times (see Table 1) and procedures (as registered on videotape) was helpful in clarifying how students obtained their answers and in understanding the effect of the experimental manipulations.

First of all, there seems to be a clear relationship between the duration of a solution process and its overall quality. As Table 1 shows, students who gave the linear answer required about one minute less than students who solved the problem correctly, and this difference was found in all three interview conditions. This is not surprising considering our previous research findings: the linear model is self-evident for many students and affects their thinking in an immediate and spontaneous way, whereas they experience the quadratic relation between lengths and area often experienced as counter-intuitive (De Bock et al., 2002a). A quick resposnse to a non-linear problem is often an indication that a student is (mis)led by linear thinking.

More importantly, the analysis of solution times and procedures revealed substantial differences between the interview conditions. Both in the D-condition and in the P-condition, many students did find the correct answer, but the way in which this answer was achieved differed. In the D-condition, the 16 students who found the correct answer needed relatively much time (on average 139 seconds), and they had to rely extensively on the drawing. Often, the idea to work on the drawing came up rather late in the solution process. When students gave the correct answer, many of them were still not very convinced about its correctness. Often, they indicated that reading, interpretation or calculation errors might have occurred or that they might have been overlooking a critical aspect in the problem situation. As an example, we quote from the interview with Deborah (D-condition, solution time of 194 seconds):

Deborah: [Silence of about 60 seconds. Reads the problem several times again.] "So … I think I should see … [Measures sides of small and large square.] The small floor is 12 cm and the large 36 cm. So that's 3 times. Eh … 3 times 4 tiles is 12 tiles. No, it's 3 times here and 3 times there, that's 9 times. You need 9 tiles… Wait, let me read it again [Reads the problem and thinks for a long time.] I don't know how I should calculate it. Maybe here … there's 4 tiles, and … yes! [Draws 9 of the small floors in the large floor.] 9 times more, so I take 9 times those 4 tiles. 36, you need 36 tiles."

Interviewer: "36 is your answer. Can you tell me how certain you are that that is the correct answer?"

Deborah: [Chooses 'Probably correct'] "I'm never sure about myself. I don't trust it, maybe there's something wrong with the drawing. It could be a tricky question."

Such a process contrasts with many solution procedures from the P-condition. Here, students needed on average only 76 seconds to respond correctly. In most cases, the students immediately and spontaneously started to manipulate the materials to find the solution (figuring out rather quickly that 6×6 tiles fit on the large floor, or that the small floor fits 3×3 times on the large one). Remarkably, three students in the P-condition gave the correct answer almost immediately. Once the problem situation

was explained to them, they did not require any additional time for thinking or material manipulating at all. They just "saw" the correct solution at a glimpse. Generally, students in the P-condition were moreover very convinced about the correctness of their answer. Consider for example the following fragment from the interview with Marlies (P-condition, solution time of 34 seconds):

Marlies: [Takes the small floor and fits it several times on the large floor.] "It's 9 times this small one, which has 4 tiles, so 36 tiles." [Goes immediately to fetch 36 tiles.]

Interviewer: "So you brought 36 tiles. Wait a moment before putting them on the floor. First, can you tell me how certain you are that that is the correct answer?"

Marlies: [Chooses 'Certainly correct'] "It is correct. I just showed you that it's 9 times more. Why would I need to doubt about it? I am just sure"

Posttest

Table 1 also contains the results on the non-linear posttest item. It aimed at testing whether students who had solved the non-linear problem correctly during the interview would do this on the posttest as well (taken one or two days later). Apparently, this was not really the case. The S-condition student who found the correct solution solved the posttest problem correctly as well (again, by making a drawing). But in the other two conditions hardly any effect of the interview experience was found: While 16 of the 24 students in the D-condition interview profited from the drawings to find the correct answer, all of them again reasoned linearly on the posttest. And whereas 20 of the 24 students found the correct answer in the P-condition interview, only two of them solved the posttest item correctly (and all others, except one, again made a linear error).

Finally, of the four students from the P-group who failed to solve the problem by themselves but who could act out and see the correct answer at the end of the interview task, only one student solved the non-linear posttest item correctly; the other three gave the same answer on the posttest as during the interview.

CONCLUSIONS AND DISCUSSION

In previous research, students' overreliance on linearity was often observed by means of tests containing traditional, scholastic word problems. The current study has shown that this has an important impact on students' solution behavior. When students who made a linear error on the pretest were involved in an interview with more meaningful, authentic performance tasks, they approached the problems very differently, and they were less tended to overgeneralise linear methods. As such, our results confirm those observed by other scholars and for other kinds of modelling problems (for an overview, see Verschaffel et al., 2000).

More importantly, our study has additionally shown that offering meaningful, authentic performance tasks affected students' problem solving behavior only at that moment itself. At a posttest (again with traditional word problems) taken shortly afterwards, nearly all students again gave a linear answer to the non-linear problem. By means of more fine-grained research, focusing on the differences in the

mathematical concepts, heuristic and metacognitive strategies, beliefs, assumptions, etc. that students activate when solving meaningful, authentic performance tasks versus traditional word problems, we hope to get deeper insight into the reasons why performance tasks have such a strong but at the same time such a context-specific impact on students' performance.

References

Cobb, P., Yackel, E., & McClain, K. (2000). *Symbolizing and communicating in mathematics classrooms*. Mahwah, NJ: Lawrence Erlbaum Associates.

De Bock, D., Van Dooren, W., Janssens, D., & Verschaffel, L. (2002a). Improper use of linear reasoning: An in-depth study of the nature and the irresistibility of secondary school students' errors. *Educational Studies in Mathematics, 50*, 311–334.

De Bock, D., Verschaffel, L., & Janssens, D. (1998). The predominance of the linear model in secondary school students' solutions of word problems involving length and area of similar plane figures. *Educational Studies in Mathematics, 35*, 65–83.

De Bock, D., Verschaffel, L., & Janssens, D. (2002b). The effects of different problem presentations and formulations on the illusion of linearity in secondary school students. *Mathematical Thinking and Learning, 4*, 65–89.

DeFranco, T. C., & Curcio, F. R. (1997). A division problem with a remainder embedded across two contexts : Children's solutions in restrictive versus real-world settings. *Focus on Learning Problems in Mathematics, 19*(2), 58–72.

Esteley, C., Villarreal, M., & Alagia, H. (2004). Extending linear models to non-linear contexts: An in-depth study about two university students' mathematical productions. In M. J. Høines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 343–350). Bergen, Norway.

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht: Reidel.

Greer, B. (1993). The mathematical modelling perspective on wor(l)d problems. *Journal of Mathematical Behavior, 12*, 239–250.

Modestou, M., Gagatsis, A., & Pitta-Pantazi, D. (2004). Students' improper proportional reasoning: The case of area and volume of rectangular figures. In M. J. Høines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 345–352). Bergen, Norway.

Nunes, T., Schliemann, A., & Carraher, D. (1993). *Street mathematics and school mathematics*. Cambridge, UK: Cambridge University Press.

Reusser, K., & Stebler, R. (1997). Realistic mathematical modelling through the solving of performance tasks. *Paper presented at the Seventh European Conference on Learning and Instruction*, Athens, Greece.

Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D., & Verschaffel, L. (2003). The illusion of linearity: Expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics, 53*, 113–138.

Van Dooren, W., De Bock, D., Hessels, A., Janssens, D., & Verschaffel, L. (2005). Not everything is proportional: Effects of age and problem type on propensities for overgeneralization. *Cognition and Instruction, 23*(1), 57–86.

Verschaffel, L., Greer, B., De Corte, E. (2000). *Making sense of word problems*. Lisse, The Netherlands: Swets & Zeitlinger.