Students' understanding of linear and non-linear functions: Two studies on the mediating role of external representations

Dirk De Bock
Wim Van Dooren
Lieven Verschaffel

HUB RESEARCH PAPERS 2012/40
ECONOMICS & MANAGEMENT
NOVEMBER 2012

ASSOCIATIE K.U.LEUVEN

HUB

Students' understanding of linear and non-linear functions: Two studies on the mediating role of external representations

Dirk De Bock [1 2], Wim Van Dooren [2] and Lieven Verschaffel [2]

[1] Hogeschool-Universiteit Brussel, Brussels, Belgium

[2] Katholieke Universiteit Leuven, Leuven, Belgium

Author for correspondence:

Dirk De Bock, Faculty of Economics and Management, Hogeschool-Universiteit Brussel, Stormstraat 2, B-1000 Brussels, Belgium, dirk.debock@hubrussel.be, Phone +3216604250, fax +3216604554.

Abstract

We investigated students' understanding of linear and different types of non-linear functions, and the way this understanding is mediated by various external representations. In a first study, we focus on students' ability to model textual descriptions of situations with different kinds of representations of linear and non-linear functions. Results highlight that students confuse linear and non-linear models, and that the representational mode has a strong impact on this confusion: Correct reasoning about a situation with one mathematical model can be facilitated in a particular representation, while the same representation is misleading for situations with another model. In a second study, we investigate students' ability to link representations of linear and non-linear functions to other representations of the same functions. Results indicate that students make most errors for decreasing functions. The number and nature of the errors also strongly depends on the kind of representational connection to be made. Both studies provide evidence for the strong impact of representations in students' thinking about linear and different types of non-linear functions.

Key words: affine model, external representations, inverse linear model, linearity, multiple representations, non-linearity

Mathematics educators very often emphasize the (stimulating) role of multiple external representations in mathematics. As Matteson (2006) explains, learning mathematics is like learning a foreign language. External representations are key elements in the vocabulary of that language, and students need to become fluent in their use if they want to succeed in expressing and understanding mathematical ideas with correctness and precision. Vergnaud (1997) argues that external representations are inherent to the discipline of mathematics, since a characteristic of mathematical concepts is that they can *only* be communicated through their external representations. The use of multiple external representations has also been shown to facilitate mathematical problem solving (e.g. Duval, 2002; Even, 1998; Gagatsis & Shiakalli, 2004; Kaput, 1992; Yerushalmy, 2006). The NCTM Standards (1989) therefore hold a strong plea for establishing "mathematical connections" through the use of multiple external representations:

> *Different representations of problems serve as different lenses through which students interpret the problems and the solutions. If students are to become mathematically powerful, they must be flexible enough to approach situations in a variety of ways and recognize the relationships among different points of view* (p. 84).

However, research has shown that students are not always sufficiently fluent in using external representations in the sense that they do not have the necessary diagrammatic knowledge to interact with the representations (de Jong et al., 1998), to interpret representations by linking them with reality (Ainsworth, Bibby & Wood, 1998), or to translate and switch between representations within the same domain (Even, 1998). Another reason why students do not always benefit from using multiple external representations in problem solving is that they are unable to make flexible representational choices, in terms of characteristics of the to-be-solved task and/or in terms of subject characteristics (Acevedo

Nistal, Van Dooren, Clarebout, Elen & Verschaffel, 2009, 2010; Acevedo Nistal, Van Dooren & Verschaffel, 2012, in press).

Research on representational fluency and flexibility often focuses on the concept of function (e.g. DeMarois & Tall, 1999; Leinhardt, Zaslavsky & Stein, 1990), because functions are a prominent example in the domain of mathematics where different types of representations (such as graphs, formulas, and tables) can be used (Even, 1998). More specifically, most studies focus on *linear* functions. A main reason is that previous research has shown that students of various ages often exhibit a very limited understanding of linearity and tend to assume it in situations that are not linear at all (De Bock, Van Dooren, Janssens & Verschaffel, 2007; Van Dooren, De Bock, Janssens & Verschaffel, 2008). For example, research has shown that a large majority of 10- to 12-year-old pupils answer 170 seconds to the word problem "John's best time to run 100 m is 17 seconds. How long will it take him to run 1 km?" (Verschaffel, De Corte & Lasure, 1994; Verschaffel, Greer & De Corte, 2000), although the reality referred to in this problem allows no single, precise answer. Another example was documented by Van Dooren, De Bock, Depaepe, Janssens and Verschaffel (2003): Many upper secondary students respond proportionally ($2 \times 1/6 = 2/6$) to probabilistic problems such as "The probability of getting a six in one roll with a die is 1/6. What is the probability of getting at least one six in two rolls?" Students' overreliance on linear models has already been studied extensively in a variety of mathematical domains (e.g., elementary arithmetic, algebra, (pre)calculus, probability, and geometry) and, more recently, also in physics (De Bock, Van Dooren & Verschaffel, 2011). De Bock et al. (2007) explain this phenomenon by referring to (1) the intuitive, heuristic nature of the linear model for students, (2) students' experiences in the mathematics classroom and their beliefs toward mathematical modeling and problem solving, and (3) elements related to the mathematical particularities of the problem situation in which the linear error occurs.

Some studies have pointed to the role of external representations in students' overreliance on linearity. In the domain of functions, for instance, the straight line graph prototype proved to be very appealing for many students. Leinhardt et al. (1990) mentioned several studies showing that students of different ages have a strong tendency to produce a linear pattern through the origin when asked to graph non-linear situations, such as the growth in the height of a person from birth to the age of 30. Another example in this domain is the study by Markovits, Eylon and Bruckheimer (1986). When asked 14- to 15-year-old students to draw a graph of a function that passes through two given points, students typically drew straight lines. Similarly, Karplus (1979) found that when students interpolate between two graphed data points in a science experiment, they strongly tend to connect the points using a straight line. Although, as exemplified above, several studies referred to the influence of external representations on students' overreliance on linearity, we are not aware of empirical studies focussing on the role of external representations on this well-known phenomenon. Consequently, the representational aspect remained a blind spot in the literature on students' overreliance on linearity.

In this article, we aim at setting a first step in unravelling the largely neglected relation between students' (lack of) mastery of external representations and their overreliance on linearity. Therefore we conducted two empirical studies. In the first study, which fitted into the research tradition of students' overreliance on linearity, we focused on the modeling aspect of functions. We investigated how accurate students are in connecting descriptions of realistic situations to various external representations of linear[i] and different types of non-linear functions. Because we hypothesized a mediating role of these external representations in students' tendency to inappropriately connect non-linear situations to linear models, we also investigated in a second study how well these different representations are understood. So, in this second study, functions were no longer used as models of situations but as

mathematical objects per se. More concretely, in the second study we investigated how accurate students are in connecting various representations of linear and non-linear functions to other representations[ii] of the same functions, without any contextualization of these functions.

Study 1: Connecting representations of functions to realistic situations

Aims and rationale

In this study we focused on students' modeling competencies with linear and non-linear functions. A key – but not at all obvious – step in a modeling process is the translation of a problems situation in a mathematical model. Students' difficulties in selecting a model and the link with the representational mode in which that model was given, were, for instance, addressed by Frejd and Ärlebäck (2011) who investigated Swedish secondary students' mathematical modeling competencies. Selecting a model was one of the seven sub-competencies these researchers focused on. To measure this (and the other) modeling sub-competencies, a slightly adapted version of a research instrument developed by Haines, Crouch and Davis (2000) was used. In two items of their instrument a "realistic" situation was described and students had to link this situation to an appropriate mathematical model to be chosen from five given alternatives. Models were either represented in a graphical or in a formula representation. The results indicated that selecting a model was one of the sub-competences students showed least proficiency, both in a graphical or formula representational mode. Also the previously mentioned article by Leinhardt et al. (1990) already made a link between students' modeling competence and their preference for linear patterns.

Because so far, the relation between students' modeling competence, their overreliance on linearity and their (lack of) mastery of representations was never systematically investigated, both in terms of models and in terms of accompanying representations, we set up this first study. More concretely, we investigated: (1) How accurate are students in connecting descriptions of realistic situations to linear and different types of non-linear models, and (2) Do accuracy and model confusion depend on the representational mode in which a model is given? Because we expected that model confusion would be more likely to occur with models that share at least some characteristics with the linear model and we anyway had to make a selection in the infinite domain of non-linear models, we worked in this study with three specific types of non-linear models that are conceptually most related to the linear model ($y = ax$): We worked with inverse linear models ($y = a/x$), affine models with positive slope ($y = ax + b$ with $a > 0$ and $b \neq 0$), and affine models with negative slope ($y = ax + b$ with $a < 0$ and $b \neq 0$). Arguably, all these models share some characteristics with the linear model but not all. Positive affine functions for instance share with linear functions the property that their graph has the shape of a straight line, and that the same increase Δx in x always results in the same increase Δy in y. But while in linear functions doubling x implies doubling y, this does not hold for affine functions. In the rest of this article, we will use the collective term "almost linear" functions for these three types of non-linear functions. Moreover, also because realistic situations dealing with negative x- or y-values are not always easy to find, we limited ourselves to realistic situations that we could describe with a model that is situated in the first quadrant.

Method

Sixty-five students in the first year of Educational Studies of the University of Leuven participated. These students had successfully finished secondary education and typically also

three years of non-university higher education. Although they all followed the obligatory mathematics courses in secondary school, this was in most cases not the core of their curriculum. In these mathematics courses, solving realistic problems and the applicability of basic functions such as the ones central in our study receive quite some attention.

Participants were confronted with a written multiple-choice test consisting of twelve descriptions of realistic situations they had to connect with an appropriate mathematical model. For each situation the appropriate model was either linear or "almost linear": Inverse linear, affine with positive slope, or affine with negative slope. These models were given either in graphical, tabular or formula form (each representation was provided in one third of the cases). Three examples of descriptions of situations to be connected to the appropriate model in one of the three representations as used in the study are listed in Figure 1. Situations were chosen so that there was a strong and clear fit with one of the provided models. We are aware that models never perfectly fit to a realistic situation, but the model we considered as the correct one for a situation provided an unquestionably better fit than the other three models.

Insert Figure 1 about here

Responses of 64 participants[iii] were statistically analyzed by means of the generalized estimating of equations (GEE) approach within SPSS (Liang & Zeger, 1996). This procedure allows to analyze repeated (and thus possibly correlated) categorical observations within series of individuals, and to appropriately correct for the correlation between measurements in order to make inferences. Given the dichotomous nature of the dependent variable (i.e., is a particular response alternative chosen or not), a logistic

regression (modeling the probability that a correct response is given, depending on the type of function and the representational mode as explanatory variables) was appropriate.

Results

Insert Table 1 about here

Table 1 shows the percentage of correct assignments for the different models underlying the verbal descriptions and for the different representational modes The logistic regression analysis first of all revealed a main effect of model, *Wald Chisquare* (3), = 38.472, $p <$ .00015, indicating that students' accuracy in connecting situations to the appropriate mathematical model depended on the type of model involved: The percentage of correct matches for an underlying linear model was significantly higher than for an inverse linear, affine with positive slope, and affine with negative slope model. The analysis did not reveal a main effect of representation, *Wald Chisquare* (2), = 5.297, $p = .071$: The percentages of correct assignments for the three representational modes did not differ significantly. Finally but most importantly, an interaction effect between model and representation was found, *Wald Chisquare* (6), = 45.111, $p < .0015$: The percentage of correct matches for a given underlying model depended on the representational mode in which that model was given, and for different underlying models, different representational modes led to higher percentages of correct matches. Although for an underlying linear model, all representations were quite good, the percentage of correct assignments in the formula mode was significantly lower than in the tabular ($p = .049$) and graphical mode ($p = .028$). For the inverse linear model, the result for formula was significantly better than for table ($p < .0015$) and graph ($p < .0015$). Underlying affine models with positive slope elicited significantly more correct matches in

the graph mode than in the two other representational modes (table $p < .0015$, formula $p = .014$) , while underlying affine models with negative slope elicited significantly less correct matches in the formula mode than for the two other modes (table $p < .0015$, graph $p < .0015$).

These results indicate that the graph was the best representation in all cases, except for the inverse linear relationships. For these relations the formula proved to be more supportive. A possible explanation is that students tend to associate characteristics of linear graphs (straight lines) and linear tables (equal distances) with graphs and tables in general. Absence of these prototypical characteristics in graphs and tables of inverse linear relationships might have refrained students from choosing these representations, leading to a decrease of correct matches for the inverse linear model when given in graphical or tabular mode. Another observation is that formulas seem to be misleading for affine relations. This might have been a side effect of the fact that the situations to be modeled were given in a "$y = b \pm ax$" form while the formulas to be matched with were given in a "$y = ax + b$" form.

Based on the number of correct responses, we conclude that students can interpret all representations (since correct matches for all representations varied between 83% and 74% and mutual differences were not significant). Students are also able to identify underlying mathematical models (more than 80% of the students detected the underlying model in at least one of the three representations), but some representations support one underlying model better than other underlying models, while others put students on the wrong track. To obtain a better understanding of these findings, an (quantitative) error analysis was conducted.

Insert Table 2 about here

Table 2 shows the percentage of erroneous assignments for the different models in the distinct representational modes. The error analysis revealed that in situations for which the

underlying model was inverse linear, (direct) linear errors were most frequently made, especially in the tabular and graphical representational mode. This finding confirms results of previous studies on students' overreliance on the linear model (De Bock et al., 2007) when an "almost linear" model is appropriate. Students are most likely misled by characteristics of representations of the linear model (i.e., equal distances and straight lines), which are prominently present in respectively tabular and graphical representations.

Also in situations in which the underlying model was affine with positive slope, linear errors were most frequent, but here these errors mainly occurred in the formula and tabular representational mode. Apparently, for this model, which comes closest to the linear model, students *see* the value at 0 or the Y-intercept in the graphical representation, while this element seems to be more often overlooked in the other representations.

Finally, in situations in which the underlying model was affine with negative slope, inverse linear errors were most frequent, especially in the formula representational mode. A straightforward explanation refers to the decreasing nature of both models. Apparently, for many students, the independent variable in the denominator is more salient than the negative sign in the numerator. Another explanatory element refers to the attraction of the "doubling/halving" prototype in situations of decrease, which is most prominent in the formula representation. Students likely experience this prototype both in daily life and from school word problems that are typically found in textbooks for teaching inverse proportionality such as "For 4 painters, it takes 12 days to paint a bridge. How many days does it take 8 painters to do that job?" (Van Dooren, De Bock, Hessels, Janssens & Verschaffel, 2005) and tend to over-generalize it to situations of decrease which are not inverse proportional.

Conclusions

Our results show that students are very proficient in relating descriptions of realistic situations to models when the situation described is a linear one. In cases where the situation is "almost linear" (i.e., affine with a positive or negative slope, or inverse linear), there is, however, a strong tendency to fall into the linearity trap and to connect the situation also to linear models (and for inverse linear situations to some extent also to affine models with a negative slope). These results are in line with several other studies showing the "default" role of the linear model, here in cases where an "almost linear" model is appropriate (inverse linear, affine with positive slope, and affine with negative slope models) (De Bock et al., 2007; Van Dooren et al., 2008).

Results also indicate that the representational mode has a strong impact on students' modeling accuracy and on their tendency to inappropriately connect non-linear situations to linear models. This last tendency was most clear for situations having an underlying model that was inverse linear or affine with positive slope, but it was always mediated by the representation in which the model was given. Apparently, a particular representation may highlight aspects of non-linearity that are easily noticed by students (e.g. the Y-intercept of a graph or the distances in a table) and therefore facilitate correct reasoning, but be misleading when representing a situation with another model.

Study 2: Connecting representations to representations

Aims and rationale

Study 1, as well as most research on improper linear reasoning so far, was related to mathematical modeling, i.e. to tasks in which real-life situations had to be expressed in mathematical terms. Although students' overreliance on linearity occurred in all representations, the representational mode proved to play a mediating role in these modeling

tasks. Much less research exists on students' (lack of) understanding of representations per se and how this might be related to their tendency toward improper linear reasoning. Students' conceptual understanding of representations in an abstract mathematical context, where thus no modeling of real-life situations needs to take place, was examined in Study 2.

More specifically, Study 2 focuses on students' fluency in linking multiple representations of functions. The importance of this skill in the mathematics curriculum is widely acknowledged (e.g. Elia, Panaoura, Gagatsis, Gravvani & Spyrou, 2006). More specifically, we hypothesize that students' limited conceptual understanding of the domain of linear functions will lead to difficulties in distinguishing them from "almost linear" functions (cf. supra). We assume that this limited conceptual understanding will interfere when students have to connect various representations of these functions to each other. Therefore, Study 2 investigates (1) how accurate students are in connecting representations of linear and "almost linear" functions to other representations of these same functions, and (2) whether accuracy in connecting representations and the confusion between linear and "almost linear" functions depend on the nature of the external representations that have to be connected to each other.

Method

The same sixty-five students from Study 1 also participated in Study 2, but the order in which they participated in both studies was counterbalanced. They were confronted with a written multiple-choice test consisting of twelve items. In each item, a graph, formula or table was provided, that had to be linked to one of four graphs (when a formula or table was given), to one of four formulas (when a graph or table was given) or to one of four tables (when a graph or formula was given). Only one of the graphs, formulas or tables accurately represented the same function as the given graph, formula or table. The test offered graphs, formulas and tables of the functions that were already used in Study 1: Linear functions and three types of

"almost linear" ones (inverse linear, affine with positive slope, and affine with negative slope). An example item is shown in Figure 2. The twelve items were offered in a random order to students. Data were analysed by means of a repeated measures logistic regression analysis followed by multiple pairwise comparisons. Additionally, an error analysis was conducted in order to investigate the most frequently chosen incorrect answering alternatives.

Insert Figure 2 about here

Results

The logistic regression first of all showed a significant main effect of the type of function, *Wald Chisquare* (3) = 28.322, $p < .0015$. Pairwise comparisons indicated that accuracy is considerably higher for items dealing with linear and positive affine functions (with average accuracy rates of 0.90 and 0.88, respectively) than for items dealing with negative affine and inverse linear functions (with average accuracy rates of 0.73 and 0.77, respectively). Thus, students had more difficulties in appropriately linking representations of functions where a larger value of x implies a smaller value of y than representations of functions where a larger value of x implies a larger value of y.

Second, the logistic regression analysis indicated a main effect of the type of representational connection that students had to make, *Wald Chisquare* (5) = 97.109, $p < .0015$. Pairwise comparisons indicated that linking a given graph to a table and linking a given table to a graph were done best by students (average accuracy of 0.95 in both cases), while linking a given formula to the correct table and linking a given table to the correct formula were significantly more difficult (average accuracy of 0.89 and 0.84, respectively). The lowest accuracies occurred for linking a graph to a formula and a formula to a graph

(accuracies of 0.76 and 0.62, respectively). This indicates that students can deal best with representations that involve concrete function values: Given concrete function values in a table, the correct graph and/or the correct formula can be rather easily retrieved (and reversely), while the link between graphs and formulas is more difficult, probably because there are no concrete function values given as an intermediate step.

Third, and most important, a significant 4 × 6 interaction effect between the type of function and the type of representational connection was found, *Wald Chisquare* (15) = 31.314, p = .006. This interaction effect globally indicates that for some functions, certain representational connections are made easier than for other functions. In order to get a good understanding of this complex interaction, the accuracy rates for the various types of items and connections between representations are summarized in Table 3. To allow a proper interpretation, the data from Table 3 are completed with data coming from an analysis of the most frequently chosen incorrect answering alternatives. As can be seen in Table 3, the items dealing with linear functions were generally solved rather well (average accuracy of 0.90), but nevertheless connecting a graph to a formula and vice versa led to errors in about 20% of the cases. A further analysis showed that when students had to select a graph given a linear formula, they often selected an affine (positive or negative) graph, probably because they were misled by the fact that these graphs are also straight lines. When selecting a formula for a given linear graph, affine formulas were hardly selected. Nearly all incorrectly selected formulas were inverse linear ones. Apparently, students were aware that they had to select a formula without an intercept, but sometimes picked a function with x in the denominator of the formula.

Insert Table 3 about here

Also the items dealing with positive affine functions were generally solved rather good (average accuracy 0.88), but choosing the right graph for an affine formula led to a considerable number of mistakes. About half of the errors were due to the fact that students selected the graph of a negative affine function instead of the positive one, indicating that they were looking for a straight line graph that does not pass the origin, but not taking into account the positive value of the slope in the formula.

The affine items with a negative slope were the most difficult ones overall (average accuracy 0.73). When looking more closely at the types of representational connections, it becomes clear that not only linking graphs to formulas led to errors (as for the other functions), but also linking tables to the correct graph, and linking tables to the correct formula. In all these cases, about half of the errors were due to the fact that students selected the positive affine alternative. Apparently, students most often recognize the affine character of a function (be it in a graph, formula, or table), but have trouble interpreting the negative slope in the formula (a negative value of a). Also in the table representation (where the negative slope can only be noticed by seeing that larger values of x imply smaller values of y) this leads to difficulties. For the negative affine functions, a second major error was noticed: In about one third of the errors, the inverse linear alternative was chosen. These students realized that larger values of x implied smaller values of y, but then erroneously selected a hyperbola graph instead of a straight (decreasing) line, or a formula with x in the denominator instead of a formula with x in the numerator but with a negative coefficient.

Finally, the inverse linear items were generally solved slightly better than the negative affine ones (average accuracy 0.77). A closer look at the error patterns shows a quite diverse picture for the different representational connections. As for the other function types, connecting the right graph to a given formula and vice versa were the most difficult tasks for students. But it is remarkable that when students had to select a graph for a given inverse

linear formula, the most frequent error (about two thirds of the errors) was the choice of the linear graph, while when students had to select a formula for a given inverse linear graph, they chose in about two thirds of the cases for the negative affine alternative. Moreover, students had also considerable difficulties in connecting the right table to a given inverse linear formula, while this phenomenon was not observed for the other type of functions. A closer look at the errors indicates that in almost all erroneous answers, the linear table was chosen instead of the correct one.

Conclusions

The results of Study 2 indicate that students have difficulties in distinguishing linear and various kinds of "almost linear" functions (inverse linear functions, affine functions with positive slope, and affine functions with negative slope). Particularly the decreasing functions (in which a larger value of x implies a smaller value of y) are less well understood: Affine functions with a negative slope are often linked to representations of affine items with a positive slope (which share the straight line graph that does not pass the origin) or with representations of inverse linear functions (which share their decreasing character). In the same way, inverse linear functions are often linked to representations of affine functions with a negative slope, but additionally, students often link an inverse linear function to representations of linear functions.

The representational connection that was most difficult was the one between a formula and a graph, and vice versa. All representational connections wherein a table was involved were made considerably better. Our hypothetical explanation for this last finding is the absence of concrete function values when linking graphs and formulas. The exemplary function values that are given in a table allow the student to test very concretely which graph or formula fits (and similarly, a formula and graph can be concretely tested against a few

alternative tables). For the items used in our test, an expert would probably be able to immediately recognize the appropriate graph for a given formula (and vice versa) without turning to concrete values, merely by comparing the formula to the global shape of the graphs. Apparently, the students involved in this study did not have that prior knowledge to make the right connections between graphs and formulas.

General discussion

The two studies reported in this article point to the important mediating role of representations in students' understanding of linear and non-linear functions. Some representations highlight aspects of non-linearity and prevent students from falling into the linearity trap, while others seem to produce the opposite effect. Parallel in both studies is that students have most difficulties with decreasing functions and their representations. However, the results of both studies also diverge in several aspects, indicating that finding a mathematical model and understanding a functional relation per se are two quite different things. In a mathematical modeling context (Study 1), graphical representations were helpful in most cases to detect the model underlying a realistic situation, while for mutually connecting representations (Study 2), tabular representations, providing concrete function values, proved to be most supportive.

Of course, these two studies also have their weaknesses. One can, for instance, point to the somewhat artificial character of the testing setting. Instead of asking students to construct a model or a representation, they only had to select the right alternative in a multiple-choice format. Moreover, they did not have other resources, such as computers or graphing calculators, to their disposal. This is rather atypical for genuine mathematics classes in Belgium. This critique especially holds for Study 1: Mathematical modeling, including the transition between a real-life situation and a mathematical model, is certainly broader and

more sophisticated than just picking a representation of the right model from a given series of representations of possible models.

In mathematical practices in and outside school, even a given table or graph has different "parts" and the art often consists in selecting the right *part* of that table or graph, i.e. the part that adequately grasps the gist of a given situation, answers a particular question that one has in mind, or clearly highlights the key features or general shape of a given relationship. For instance, as a prototype of a quadratic relationship, nobody will ever show a piece of a parabola that does not include the top of that curve. Usually computer-supported environments are employed to give representations a dynamic character and in that way support this selection and adaptation process. That functions have *several* tables and graphs – and thus that *the* table or graph of a given function does not exist – even holds for formulas. Also a "same" formula may appear in different forms. As indicated in Study 1: Although "$y = ax + b$" and "$y = b + ax$" mathematically describe the same relationship, from a cognitive point of view they may be perceived differently by students.

A second shortcoming of our studies, which is directly related to the testing format, was the absence of process-oriented data about the criteria that students used to select models or representations. We hypothesize that students relied on properties of functions and their representations, but further research should confirm this. In that respect, it would be interesting to systematically investigate how students' conceptual understanding of linear and "almost linear" functions and their representations is mediated by such functional properties. For instance, given a representation (a table, graph, or formula) of a linear, positive affine, negative affine or inverse linear function, one could ask students to indicate whether a statement as "when x doubles, y doubles" is true. One could expect that students most easily detect the incorrectness of such statement for non-linear functions in a tabular representation because this representation allows them to readily compare given function values. By

involving the role of functional properties in the research on functions and representations, a more refined view on students' selection criteria and difficulties in distinguishing linear and "almost linear" functions could be achieved.

Notwithstanding the above-mentioned limitations, an implication that could be drawn for mathematics education is the need for drawing sufficient instructional attention to representations, to discuss strengths and weaknesses of various representational forms, to match representations with each other and to link them to realistic situations, and for explicitly discussing differences between linear and different types of "almost linear" models. Modeling tasks (Study 1) as well as decontextualized matching tasks (Study 2) can alternate in concrete learning environments and can both be applied to develop students' ability to recognize mathematical functions and their characteristics, and to fluently and flexibly apply related procedures (or strategies). In our view, both type of tasks could and should also be used as a starting point for tasks relating more authentic real-world situations to mathematical models and with reflections on this relation (Greer, 2006; Verschaffel et al., 2000). Future interventions studies could empirically validate this type of learning environments, including the effectiveness by which different representations are used to strengthen students' insight in the domain of functions and their applications.

References

Acevedo Nistal, A., Van Dooren, W., Clarebout, G., Elen, J. & Verschaffel, L. (2009). Conceptualising, investigating and stimulating representational flexibility in mathematical problem solving and learning: A critical review. *ZDM–The International Journal on Mathematics Education, 41*, 627–636.

Acevedo Nistal, A., Van Dooren, W., Clarebout, G., Elen, J. & Verschaffel, L. (2010). Representational flexibility in linear-function problems: A choice/no-choice study. In L. Verschaffel, E. De Corte, T. de Jong & J. Elen (Eds.), *Use of representations in reasoning and problem solving: Analysis and improvement* (pp. 74–93). Oxon: Routledge.

Acevedo Nistal, A., Van Dooren, W. & Verschaffel, L. (2012). What counts as a flexible representational choice? An evaluation of students' representational choices to solve linear function problems.*Instructional Science, 40*(6), 999–1019.

Acevedo Nistal, A., Van Dooren, W. & Verschaffel, L. (in press). Students' reported justifications for their representational choices in linear function problems: An interview study. *Educational Studies.*

Ainsworth, S.E., Bibby, P.A. & Wood, D.J. (1998). Analysing the costs and benefits of multi-representational learning environments. In M.W. van Someren, P. Reimann, H.P.A. Boshuizen & T. de Jong (Eds.), *Learning with multiple representations* (pp. 120–134). Amsterdam: Pergamon.

De Bock, D., Van Dooren, W., Janssens, D. & Verschaffel, L. (2007). *The illusion of linearity: From analysis to improvement* (Mathematics Education Library). New York: Springer.

De Bock, D., Van Dooren, W. & Verschaffel, L. (2011). Students' overuse of linearity: An exploration in physics. *Research in Science Education, 41*(3), 389–412.

de Jong, T., Ainsworth, S., Dobson, M., van der Hulst, A., Levonen, J., Reimann, P., et al. (1998). Acquiring knowledge in science and math: The use of multiple representations in technology based learning environments. In M.W. Van Someren, P. Reimann, H.P.A. Boshuizen & T. de Jong (Eds.), *Learning with multiple representations* (pp. 9–40). Amsterdam: Pergamon.

DeMarois P. & Tall D. (1999). Function: Organizing principle or cognitive root? In O. Zaslavsky (Ed.), *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 257–264). Haifa, Israel: PME.

Duval, R. (2002). The cognitive analysis of problems of comprehension in the learning of mathematics. *Mediterranean Journal for Research in Mathematics Education, 1*(2), 1–16.

Elia, I., Panaoura, A., Gagatsis, A., Gravvani, K. & Spyrou, P. (2006). An empirical four-dimensional model for the understanding of function. In J. Novotná, H. Moraová, M. Krátká & N. Stehlíková (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 137–142). Prague, Czech Republic: PME.

Even, R. (1998). Factors involved in linking representations of functions. *Journal of Mathematical Behaviour, 17*, 105–121.

Frejd, P. & Ärlebäck, J.B. (2011). First results from a study investigating Swedish upper secondary students' mathematical modeling competencies. In G. Kaiser, W. Blum, R. Borromeo Ferri & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modeling* (pp. 407–416). New York: Springer.

Gagatsis, A. & Shiakalli, M. (2004). Ability to translate from one representation of the concept of function to another and mathematical problem solving. *Educational Psychology, 24,* 645–657.

Greer, B. (2006). Designing for conceptual change. In J. Novotná, H. Moraová, M. Krátká & N. Stehlíková (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 175–178). Prague, Czech Republic: PME.

Haines, C., Crouch, R. & Davis, J. (2000). *Mathematical modeling skills: A Research Instrument* (Technical Report No. 55). University of Hertfordshire: Dept. of Mathematics.

Kaput, J. (1992). Technology and mathematics education. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 515–556). NY: Macmillan.

Karplus, R. (1979). Continuous functions: Students' viewpoints. *European Journal of Science Education, 1,* 397–413.

Leinhardt, G., Zaslavsky, O. & Stein, M.K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60*, 1–64.

Liang, K.-Y. & Zeger, S.L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 13–22.

Markovits, Z., Eylon, B.-S. & Bruckheimer, M. (1986). Functions today and yesterday. *For the Learning of Mathematics*, *6*(2), 18–24, 28.

Matteson, S.M. (2006). Mathematical literacy and standardized mathematical assessments. *Reading Psychology, 27*, 205–233.

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

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.

Van Dooren, W., De Bock, D., Janssens, D. & Verschaffel, L. (2008). The linear imperative: An inventory and conceptual analysis of students' over-use of linearity. *Journal for Research in Mathematics Education, 39*(3), 311–342.

Vergnaud, G. (1997). The nature of mathematical concepts. In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 1–28). Hove: Psychology Press.

Verschaffel, L., De Corte, E., de Jong, T. & Elen, J. (Eds.) (2010). *Use of representations in reasoning and problem solving: Analysis and improvement* (New perspectives on Learning and Instruction Series). Oxon: Routledge.

Verschaffel, L., De Corte, E. & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction, 4,* 273–294.

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

Yerushalmy, M. (2006). Slower algebra students meet faster tools: Solving algebra word problems with graphing software. *Journal for Research in Mathematics Education, 37,* 356–387.

Acknowledgement

This research was partially supported by grant GOA 2012/010 “Number sense: Analysis and improvement” from the Research Fund K.U.Leuven, Belgium.

Table 1

Percentage of correct assignments for the four models underlying the verbal descriptions and for the three representational modes

	Graph	Table	Formula	Average
Linear	98	95	89	94
Inverse linear	70	69	92	77
Affine with positive slope	81	55	66	67
Affine with negative slope	81	81	48	70
Average	83	75	74	77

Table 2

Percentage of erroneous assignments for the linear (L), the inverse linear (IL), the affine with positive slope (A+), and the affine with negative slope (A–) model in the different representational modes (graph, table, and formula)

	Graph			Table			Formula			Total		
L	IL	A+	A–	IL	A+	A–	IL	A+	A–	IL	A+	A–
	0	2	0	2	3	0	9	0	2	4	2	1
IL	L	A+	A–	L	A+	A–	L	A+	A–	L	A+	A–
	14	3	11	27	2	3	5	3	0	15	3	5
A+	L	IL	A–	L	IL	A–	L	IL	A–	L	IL	A–
	17	2	0	41	0	2	31	2	0	30	1	1
A–	L	IL	A+	L	IL	A+	L	IL	A+	L	IL	A+
	8	11	0	6	11	2	19	23	3	11	15	2

Table 3

Overview of accuracies for different function types and representational connections

		Linear functions		
			To …	
		Graph	Formula	Table
From …	Graph		0.82	0.97
	Formula	0.80		0.94
	Table	0.97	0.92	
		Affine functions positive slope		
			To …	
		Graph	Formula	Table
From …	Graph		0.88	0.97
	Formula	0.72		0.91
	Table	0.97	0.85	
		Affine functions negative slope		
			To …	
		Graph	Formula	Table
From …	Graph		0.60	0.91
	Formula	0.43		0.91
	Table	0.83	0.71	
		Inverse linear functions		
			To …	
		Graph	Formula	Table
From …	Graph		0.66	0.95
	Formula	0.49		0.75
	Table	0.95	0.82	

Figure 1

Example items from the multiple-choice test of Study 1

Example 1 (inverse linear situation/alternative formulas)

During the war, butter was rationed. Each week, butter was delivered and fairly distributed among the people. Which formula properly represents the relation between the number of people who wants butter and the amount of butter everybody receives?

- ❑ $y = 150\ x$
- ❑ $y = 150/x$
- ❑ $y = 150\ x + 30$
- ❑ $y = -150\ x + 30$

Example 2 (linear situation/alternative tables)

Jennifer buys minced meat at the butcher's shop. Which table properly represents the relation between the amount of minced meat that Jennifer buys and the price she has to pay?

❑

x	*y*
0	8
1	-4
2	-16
3	-28
4	-40

❑

x	*y*
0	0
1	12
2	24
3	36
4	48

❑

x	*y*
0	
1	12
2	6
3	4
4	3

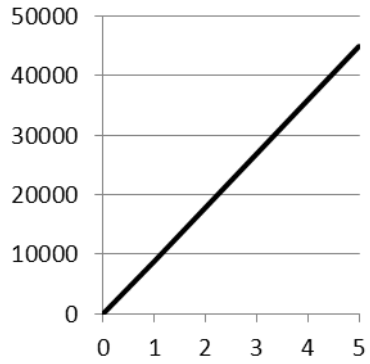
❑

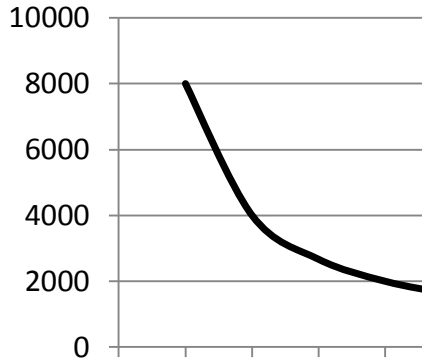
x	*y*
0	8
1	20
2	32
3	44
4	56

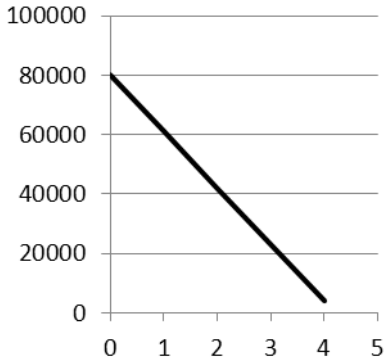
Example 3 (affine situation with negative slope/alternative graphs)

A chemical concern has a big cistern with hydrochloric. This morning they started to pump with a constant pace all hydrochloric out of this cistern. Which graph properly represents

the relation between the time elapsed and the amount of hydrochloric that is still in the cistern?

❑ 

❑ 

❑ 

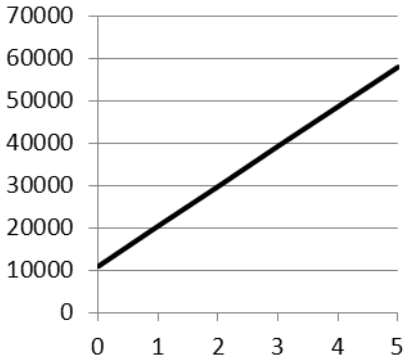
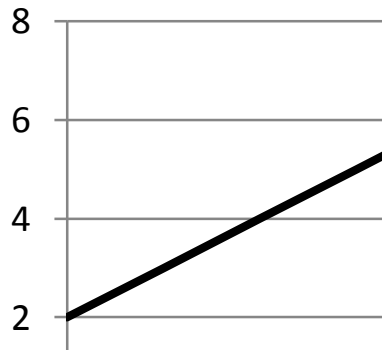
❑ 

Figure 2

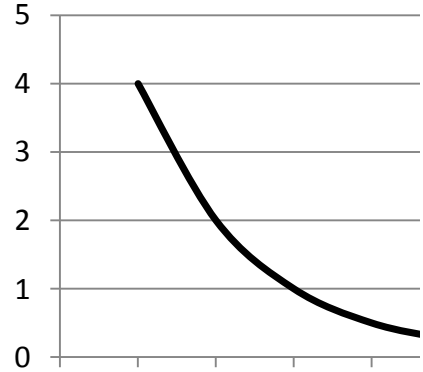
Example item from the multiple-choice test of Study 2

Choose among the four graphs below the one that describes the same mathematical relationship as the formula y = .08 x

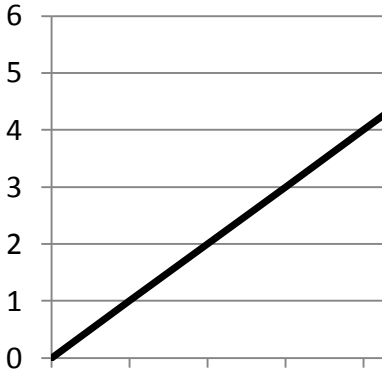
❑



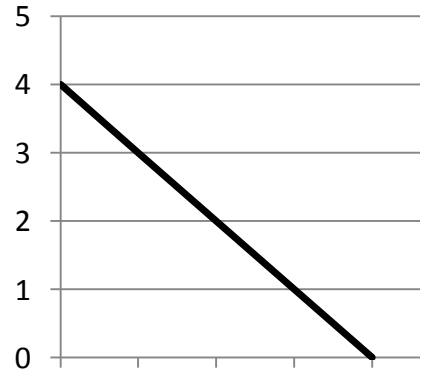
❑



❑



❑



Endnotes

[i] In this paper we use the term linear to determine functions that can be described by a formula of the form "$y = ax$". Functions that can be described by a formula of the form "$y = ax + b$" with $b \neq 0$ are labelled as affine functions.

[ii] In the rest of this paper, we will use the term "representation" to denote "external representation".

[iii] Responses of one participant were dropped from the analysis because they were incomplete.