Development of a modelling learning path

van Buuren, O.P.M.

Citation for published version (APA):
Learning to use and create formulas for constructing computer models*

Abstract

A learning path has been developed on system dynamics based graphical modelling, integrated into the Dutch secondary physics curriculum, starting from the initial phases (age: 13-15 years). In this paper, the focus is on students’ conceptions of formula and variable, on their ability to connect graphical stock-flow diagrams to difference equations, and on their ability to construct simple formulas independently. The design of a learning path on formulas and variables as part of the modelling learning path and the research findings of field testing of this part of the learning path with lower secondary students are presented.

The main conclusions are the following. The introduction of operational definitions of formula and variable and the use of formulas in educational software contribute to a clearer notion of formula and variable, as is required for modelling. Part of the students have problems translating graphical stock-flow diagrams into difference equations and vice versa. The fact that the integration time step from the equations is not visible in graphical stock-flow diagrams seems to be related to these problems. Results with respect to students’ abilities to construct formulas are promising, but strategies that students can use for constructing formulas need to be enhanced.

6.1 Introduction

In the past decades, attention for computational modelling in education has grown. In the Netherlands, for example, computational modelling has recently become a more important part of the physics curricula for upper secondary education (Commissie Vernieuwing Natuurkundeonderwijs havo/vwo, 2010).  

* This chapter has been submitted for publication as Van Buuren, O., Heck, A., & Ellermeijer, T. (2013). Learning to use and to create formulas for constructing computer models in lower secondary physics education.
Although positive effects of modelling have been reported in research literature (e.g., Doerr, 1996; Van Borkulo, 2009), it also is reported that educational results of modelling activities do not always meet expectations (cf., Doerr, 1996; Schecker, 2005). Several authors report on difficulties that students may have when designing or adapting models (Bliss, 1994; Kurtz dos Santos & Ogborn, 1994; Sins, Savelbergh, & Van Joolingen, 2005; Lane, 2008; Westra, 2008; Van Borkulo, 2009; Ormel, 2010) or when interpreting model output (Cronin, Gonzalez, & Sterman, 2009). Doerr (1996), Lane (2008), and Savelbergh et al. (2008) state that more research is needed for successful integration of modelling in education.

Because a concrete implementation of a complete, well-designed and thoroughly tested learning path on modelling into the physics curriculum does not exist yet in the Netherlands, in 2008 we started to develop a learning path on quantitative computational modelling. For several reasons, we decided to start this learning path from the initial phases of physics education (age: 13-14 years) and to integrate it into the physics curriculum (Van Buuren, Uylings, & Ellermeijer, 2010). The first two years of this modelling learning path have now been designed and tested in school practice in an iterative research and development approach (Van Buuren, Heck, & Ellermeijer, 2013a).

![Figure 6.1](image)

**Figure 6.1**: Difference equation consisting of one stock and two flow variables, and its graphical representation. The independent variable $t$ and the integration step size $\Delta t$ are not visualised.

The modelling approach used is the graphical system dynamics approach developed by J.W. Forrester (Forrester, 1968). From a mathematical point of view, this mainly concerns the numerical solving of (systems of) one-dimensional difference (or differential) equations. The modelling tool of Coach 6 (Heck, Kedzierska, & Ellermeijer, 2009; Heck, 2012) is used as the computer learning environment for this approach. In this graphical approach, difference equations are not entered in the model as formulas. Instead, they are represented by means of stock-flow diagrams: combinations of stock icons, represented by rectangles, and flow icons, represented by thick arrows (Figure 6.1). Variables (or constants) that are not explicitly part of a difference equation are referred to as auxiliary variables and are represented by circular icons. For brevity, we refer to variables represented by flow icons as flow variables or flows, and variables represented by stock symbols as stock variables or stocks. It must be noticed, however, that the type of a variable is not an intrinsic property, but can be manifold. It depends on the roles of the variable in the equations. Henceforth,
we refer to this approach as ‘graphical modelling’. A more detailed description of graphical modelling is given by (Van Buuren, Uylings, & Ellermeijer, 2011).

Although the difference equations are not entered explicitly as mathematical formulas in this graphical approach, a certain level of understanding of the separate roles of the various variables in the equations is still required; at least students must determine by which type of icon a variable must be represented in a stock-flow diagram. Many researchers (Doerr, 1996; Hogan & Thomas, 2001; Schecker, 2005; Westra, 2008; Van Buuren et al., 2011) have found that even upper secondary students do not automatically possess the required levels of understanding. In addition, students still need the ability to use formulas for variables that are defined by direct relations instead of difference equations. In the pilot phase of the research study, we noticed that many students at the beginning of the third year of secondary education (age: 14-15 years) have difficulties using direct relations in graphical computer models (Van Buuren, Uylings, & Ellermeijer, 2012). When asked to define a flow variable in a graphical model by means of a formula, a significant number of students tried to define this variable by means of numbers or by means of expressions consisting of numbers only, instead of formulas consisting of symbols. A post-test at school revealed that 59% of 80 students understood the calculation process for the model presented to them on a numerical level, but only 16% of the students were able to reconstruct the correct direct relation that corresponds to part of this process. A questionnaire revealed that the notion of formula is unclear and varies among lower secondary students. An expression consisting only of numbers was considered as a formula by approximately two thirds of all students. The same questionnaire, given to students of 23 classes at 7 other Dutch schools, revealed that such unclear notions exist at these other schools as well and even with first year upper secondary students. Similar problems with expressions have been found by Van Buuren et al. (2011) in an earlier pilot project with students at the end of the third year of secondary education.

In the intervention described by Van Buuren et al. (2012), students not only had to use a formula in a computer model, they also had to (re)construct it. Construction of formulas appears to be difficult, even for higher level students (see for example Clement, Lochhead, & Monk, 1981; Tall & Thomas, 1991; MacGregor & Stacey, 1993; Westra, 2008, p.206; Schaar, Vos, Ellermeijer, & Goedhart, 2011). This holds even more for the construction of difference equations (Verhoef, Zwarteveen-Roosenbrand, Van Joolingen, & Pieters, 2013). Reconstruction of formulas apparently is difficult too. But in order to be able to understand the calculation process for a model, students must be able to construct or reconstruct the correct formula.

---

2 By a direct relation, we mean a mathematical relationship between symbolized quantities in which at least one quantity can be isolated and written as a closed form expression of the other quantities.

3 With ‘numerical level’ we mean: using numbers, not symbols and formulas.

4 With ‘reconstruction of a formula’, we mean the construction of a formula that has been derived earlier on for or by the students. Students are not stimulated to learn this formula by heart.
design computer models themselves, eventually students must learn to construct formulas. Therefore, explicit attention must be paid to reconstruction and construction of formulas. Within mathematics education, this process of construction of formulas usually is called mathematization. Uhden, Karam, Pietrocola, and Pospiech (2011) argue that, for this part of the modelling process, more analysis is required with respect to differentiation between levels of mathematization and that a teaching strategy must deal with these different levels. Clement, Lochhead, and Soloway (1980) found that computer programming can be helpful for students to learn to construct formulas. This may also hold for computer modelling.

For these reasons, both in the instructional materials and in the classroom experiments we have paid much attention to students’ understanding and use of variables and formulas in graphical models and to the construction of simple formulas. We have carefully designed a learning path on variables and formulas in the context of modelling.

In this paper, we present the design of this learning path and the research findings of field testing of the instructional materials in school practice. It is organized as follows: in Section 6.2 we present the conceptual framework, in Section 6.3 the research methodology and research questions, in Section 6.4 the design of the learning path, in Section 6.5 an overview of its implementation into the physics curriculum, and in Section 6.6 the general setting and research instruments. In Sections 6.7, 6.8, and 6.9, this implementation and these research instruments are described in more detail, together with the results of field testing for the three most important parts of the learning path on variables and formulas. Finally, in Section 6.10 we reflect on the outcomes, answer the research questions and draw main conclusions.

6.2 Conceptual framework

Part of the students’ difficulties with variables and formulas is that both the terms variable and formula have various meanings in mathematics and science (Heck, 2001; Malisani & Spagnolo, 2008). For a proper conception of formula, a correct conception of variable is of great importance (Van Buuren et al., 2012, 2013a). Three uses of variable are important for our learning path:

1. Placeholder: a variable that stands for one number, known or unknown. According to Heck (2001), in this sense a variable can be regarded as a rather static entity by a student: it makes no sense to him or her to consider change of the variable.

2. Generalized number: an indeterminate number that appears in generalizations and in general methods (Malisani & Spagnolo, 2008). The variables in a physics formula, expressing a general relation between physics quantities, such as the relation between mass, volume, and density, can be considered generalized numbers.
3. Varying object: a symbol for an object with varying value (Heck, 2001), often in functional relationship to another variable, as “a thing that varies” (Malisani & Spagnolo, 2008), with a “changing nature” (Graham & Thomas, 2000). This notion of variable is essential for dynamical processes, governed by difference equations, but can also be important for direct relations, such as Ohms’ law, when the functional relationship between the variables is considered.

Usually in traditional school practice in physics, as soon as a formula is applied to a specific situation, the generalized numbers in it are turned into placeholders. In this way, at school an understanding of variable on the level of placeholders suffices for a long time. Such a limited notion of variable and the lack of a clear definition of the term formula may explain the findings of Van Buuren et al. (2012) summarized above, at least the students’ use of expressions consisting of numbers only, instead of ‘true’ formulas, in computer models.

In mathematics education, mathematical models (formulas) are often given ready-made. Strategies for constructing them are not taught in mathematics classes (Schaap, Vos, & Goedhart, 2011). Students have only few opportunities to create their own models and construct their own formulas. One example of difficulties that students can have when constructing formulas is the reversed equation error identified by Clement et al. (1981). The error is that, given a problem statement about a proportional relationship between two quantities, many students connect the coefficients in the equation to the wrong quantities. The reversed equation error and causes of this type of error are much discussed in mathematics education research (see, for instance, MacGregor & Stacey, 1993; Fisher, Borchert, & Bassok, 2010). According to Davis (1984), students use the symbols in the equation as units or labels instead of variables. The error can also be the result of trivial syntactic or other non-operative approaches (Clement et al., 1981).

The understanding of variables at the level of varying objects and the notion of a formula as (possibly) consisting of variables is in our perspective on modelling of crucial importance. Therefore, we are of opinion that the development of understanding of formulas and variables and of their connection to graphical models must be tuned carefully. The main question is how to do this. The following recommendations have been made (Van Buuren et al., 2012):

- A clear operational definition of the term formula must be given to the students.

---

5 An example that is often referred to is the students-professors problem. This problem is formulated as follows: 'At a university, there are 6 times as many students as professors. Give the relation between the number of professors \( P \) and the number of students \( S \).' The answer is often incorrectly formulated as ‘\( 1P = 6S \).’
• Students must get a notion of variable on the level of varying object. In particular, they must learn to distinguish between a constant and a variable as varying object.

• In graphical models, difference equations are not entered in their usual mathematical form into the model; instead, they are represented by stock-flow diagrams. This may give rise to the misconception that there is not a formula involved at all. The fact that for a stock variable an initial value needs to be entered into the model may contribute to the misconception that the stock is defined only by a constant number and not by a formula, as we noticed on occasion in classroom. Therefore, the relation between a difference equation and the graphical symbols must be made explicit to students.

• The construction of formulas needs more attention.

6.3 Research methodology and research questions

In this paper we describe a research project in which we try to develop a learning path on the understanding of the concepts of formula and variable at levels required for modelling, dealing with students’ difficulties identified in earlier work (Van Buuren et al., 2012) and incorporating the recommendations. We refer to this learning path as a ‘partial learning path’ because it forms part of the complete learning path on modelling (Van Buuren et al., 2013a). Our research approach can be classified as design research (Van den Akker, Gravemeijer, McKenney, & Nieveen, 2006): instructional materials are designed, tested in classroom, and redesigned in several cycles. Based on the test results, possibly a local theory of the subjects involved can be developed.

Research questions are as follows. To what extent can students
1. grasp the notion of formula as providing a relation between symbols representing variables and constants?
2. understand the difference between constants and variables?
3. utilise conceptions of variable and formula for modelling?
4. construct simple direct relations?
5. interpret stock-flow diagrams as difference equations and vice versa?
6. construct simple difference equations?

In these research questions, students must be understood as students involved in this learning path in the third year of secondary education.
6.4 Design of the partial learning path on formulas and variables

6.4.1 Design considerations

Targets of the partial learning path on formulas and variables are that students develop the abilities to use formulas in computer models, to distinguish between the building blocks of system dynamics based models, and to construct simple formulas themselves. To this end we adopt the four recommendations of Van Buuren et al. (2011) listed at the end of Section 6.2. In addition, students must learn to distinguish between difference equations and direct relations. Also, there must be a need for students to use formulas. Furthermore, the general design principles described by Van Buuren et al. (2013a) are used. These are the creation of bases of orientation before new concepts are introduced and design criteria for modelling learning activities, especially the limitation of the number of learning goals for computer tasks. This leads to the setup of the partial learning path described below.

6.4.2 Outline of the partial learning path

1. First, variables in the sense of measurable quantities and units are introduced, followed by an introduction of a few word formulas that are soon abbreviated and turned into ‘true’ formulas, consisting of symbols. This fairly traditional approach suffices for students to learn to deal with formulas containing variables on the level of placeholder.

2. Because students must learn to construct simple formulas themselves, (word) formulas are not just given to our students (as is often the case in traditional methods), but they are constructed, usually starting with comprehensible situations and calculations with simple numbers. These are subsequently generalized into (word) formulas, by replacing the numbers in the initial phases by words and subsequently by symbols. This approach may work for the first formulas, involving sufficiently concrete concepts, like mass, density and velocity. For more abstract concepts, other approaches are required, but these occur at a later stage in physics education, when students should already have grasped the notions of formula and variable.

3. In order to prepare students for the distinction between difference equations and direct relations, Δ-notation is used immediately from the first difference equation, early on in the learning path. In a pilot project, for these difference

6 Actually, this was a suggestion from students in the third year of secondary education. Confronted with a change in notation for the formula for distance, velocity, and time from \( s = v \cdot t \) to \( \Delta x = v \cdot \Delta t \), they argued that it would be easier for them to use Δ-notation from the beginning.
equations the term ‘change formulas’ was used (Van Buuren et al., 2011), but a number of students at the end of the third year of secondary education understood this term as ‘direct relations consisting of quantities that may change, instead of constants’, and not as difference equations. From this, we inferred that students probably would be able to distinguish between variables and constants, and that the more neutral term Δ-formula would be better.

4. As preparation for the notion of variable as varying quantity, students are introduced at an early stage into numerical integration. At this stage, the integrand is a rate of change that is given in graphical and tabular format as a function of time. In this way, varying quantities already appear in graphs and tables.

5. As preparation for the introduction of the notion of formula as relation between symbolized variables, the equivalence of an expression consisting of numbers (a ‘simple calculation’, such as 9.1÷4.2) and its outcome (one number) is discussed at an early stage. An understanding of this equivalence helps in clarifying the difference between a formula consisting of symbols and an expression consisting of numbers.

   Up to this point, there is no clear need or advantage for students to use formulas and to understand variables on the level of varying object. In the following three steps, this need is created, students are provided with an operational definition of formula, and further steps are made towards the notion of variable as varying quantity.

6. Just before the use of formulas becomes unavoidable, a formal operational definition of formula as relation between at least two symbolized physical quantities is introduced, together with the notions of equation and simple calculation. Students get the following instructions:

   “1. A formula provides a (computational) relation between two or more different physical quantities.

   2. These physical quantities are abbreviated, for surveyability. The abbreviations are called ‘symbols’.

   3. The values of the physical quantities are not all fixed in advance.

   4. In most computer programs, only one physical quantity can appear at the left side of an equal sign. That must be the physical quantity that is going to be calculated by the computer program.

   You, as a human being, usually start an exercise by filling in numbers into the formula. For example, \( R = 5 \, \Omega \) and \( U = 3 \, \text{V} \) can be entered into the formula \( R = \frac{U}{I} \). We get \( 5 = \frac{3}{I} \). This, we will call an equation. At this point, the value of ‘I’ actually is fixed, but this value is still unknown. An equation can be solved, that is: we can search and find the value for the unknown quan-
ity. For example, the above equation can be transformed into \( I = \frac{3}{5} \).

What’s left is a simple calculation that can be carried out easily.”

Subsequently, students are given a few exercises in which they practise with classifying expressions as formula, equation, or simple calculation, respectively. Of course, from a more strict mathematical point of view, a simple calculation can also be considered as an equation, and a formula can be considered as an equation with several unknowns. In mathematics and physics, the meanings of the terms formula and equation are often context dependent and therefore somewhat vague for novices. At this stage of students’ development we consider it important for students to be provided with clear operational definitions. This must help them to develop better notions of formula, variable, constant and number. At a later stage, these notions can and must be refined.

7. Subsequently, a need for formulas must be created. This is done by providing students with a large amount of data (from measurements) on which the same calculation must be carried out many times. This shows the advantage of using the computer learning environment. In this environment, the use of formulas becomes a necessity.

8. Simultaneously, a step is made towards the notion of variable as varying quantity. For the calculations in the computer learning environment, a direct relation is used. For each calculation, the variables can be considered as placeholders, but from one calculation to another, the variables are clearly changing. A table containing the data and the outcomes of the computer calculations may improve students’ understanding by visualizing this aspect of change.

Next, the notion of variable as varying quantity must be introduced, students must learn to construct direct relations, and they must learn to use direct relations in a graphical computer model. For this, a process of numerical integration is used. A realistic process that can be easily visualised and in which measurements can be done makes the process of numerical integration concrete for students. Doing measurements enables comparison between the computation process and the realistic situation. A discrete realistic process is preferred, because it more clearly matches the discrete, stepwise process of numerical integration. As an integrand, we need a rate of change of a varying quantity that must be defined by means of a direct relation. Students must construct this direct relation. To create a need for the notion of variable as varying quantity, this rate of change must depend on the varying quantity itself. In this case, in the iterative process of numerical integration, both the values of the quantity and the rate of change are computed from values calculated in the previous iteration. In this way, the integrand is not known beforehand (as it would be if the integrand was some previously given function of time, as in Step 4). The variable used for the rate of change cannot be considered as placeholder because its value is inevitably changing, the next value being dependent on previous values. These considerations lead us to the next steps:
9. After the introduction of the realistic process, students perform a few steps of the numerical integration process by hand, only using numbers. The varying character of the physical quantities emerges and the formal difference between variable as varying object and constant is introduced. Students are given a few exercises, showing that it may depend on the context whether a physical quantity is a variable or a constant.

10. The huge amount of calculations creates a need for a computer model and, consequently, the use of formulas. Construction of the required direct relation is stimulated by the repetitive character of the calculations. Students are guided towards this direct relation: the calculations are generalized into this relation. Several forms of the relation are also in the instructional materials, but rather hidden, and students are not stimulated to learn this relation by heart; they need to reconstruct it. Students practise the construction of simple new direct relations in a few exercises.

11. The relation between stock-flow diagrams and Δ-formulas is introduced.

12. The process of numerical integration as performed by the computer is introduced formally, taking advantage of Steps 4 and 9.

13. Students implement the direct relation from Step 10 in the graphical computer model.

14. The students’ base of orientation is extended by the addition of another modelling task for a new situation in which a similar direct relation must be constructed.

15. In the next part of the learning path, students keep practising these concepts and skills by using direct relations in computer models, interpreting graphical models consisting of networks of direct relations, by constructing simple formulas, and by distinguishing between variables and constants.

Finally, four important steps with respect to the difference equations and stock-flow diagrams are made:

16. Students must learn to understand that the type of a variable is not an intrinsic property, but depends on its role in the equation. Therefore, following a suggestion of Van Buuren et al. (2011), two Δ-formulas are presented to the students in which one variable has a twofold role: as a stock variable in the one Δ-formula and as a flow variable in the other. The graphical model must be constructed that incorporates both Δ-formulas. This must provide students with an enhanced base of orientation for the concepts of Δ-formula, stock-flow diagram, stock variable, and flow variable. Because in graphical modelling tools, a variable with such a twofold role cannot be represented by only one icon, students must learn to use dummy variables and dummy icons.

17. Students learn to adapt Δ-formulas to graphical modelling. An example of a Δ-formula that cannot be translated at once into a stock-flow diagram is Newton’s second law written as \( \Delta v = \frac{F_{\text{net}}}{m} \Delta t \). In this Δ-formula, there is a quotient \( \frac{F_{\text{net}}}{m} \) instead of one variable. By introducing the dummy variable \( a \), the
Δ-formula is split into \( \Delta v = a \cdot \Delta t \) and \( a = \frac{F_{\text{net}}}{m} \). These two formulas subsequently can be translated into a graphical model.

18. Students are asked to construct a Δ-formula independently. This is meant to support the understanding of the relation between Δ-formulas and stock-flow diagrams, but it is also a first introduction to the construction of Δ-formulas.

19. Eventually, all elements required for graphical modelling come together. Students construct a more extended graphical model, based on formulas that are already known to them.

6.5 Overview of the implementation of the instructional design

The entire modelling learning path is integrated into the lower secondary physics curriculum. This curriculum consists of twelve physics modules. In the second year of secondary education (age: 13-14 years), which is the first year of physics education, these modules are the following: 1. Doing Research; 2. Light; 3. Density; 4. Electricity; 5. Velocity; 6. Forces and Bridges; 7. Energy and Power. In the third year of secondary education (age: 14-15 years), the modules are: 1. Solid, Liquid, and Gas; 2. Resistance and Conductance; 3. The Vacuum Pump; 4. Sound; 5. Force and Movement. A detailed description of this modelling learning path is given by Van Buuren et al. (2013a). An overview of the distribution of the steps from the design described in Section 6.4.2 over these modules is given in Table 6.1. In this paper, the focus is on Steps 6 to 19 in the partial learning path on variables and formulas. All these steps take place in the third year of secondary education. Steps 6 to 8 are implemented in the module Resistance and Conductance, Steps 9 to 14 are implemented in the module The Vacuum Pump, and Steps 16 to 19 are implemented in the module Force and Movement. In the module Sound, in between The Vacuum Pump and Force and Movement, students practise with the concepts of the preceding modules (Step 15). No real new steps with respect to variables and formulas are made; the main learning target of this module is the relational structure of graphical models (Van Buuren, Heck, & Ellermeijer, 2013b). For this reason, we restrict the description of the implementation of the learning path to the modules Resistance and Conductance, The Vacuum Pump, and Force and Movement. This description can be found in Sections 6.7, 6.8, and 6.9, respectively.
<table>
<thead>
<tr>
<th>Module number and title</th>
<th>Design steps</th>
<th>Content regarding formulas and variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Second year of secondary education (age: 13-14 years)</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 1 Doing Research | 1 | - variables in the sense of measurable quantities;  
- units; |
| 3 Density | 2 | - first word formula;  
- first direct relation; |
| 5 Velocity | 3 | - first Δ-formula;  
- first use of a graphical model;  
- process of numerical integration for a graph of a varying quantity;  
- first introduction of the equivalence between an expression consisting of numbers only and its outcome, a single number; |
| 6 Forces and Bridges | | - two new direct relations; |
| 7 Energy and Power | | - second Δ-formula;  
- introduction of stock-flow diagrams for storage and flows of energy; |
| **Third year of secondary education (age: 14-15 years)** | | |
| 2 Resistance and Conductance | 6 | - two new direct relations;  
- operational definitions of formula, equation, and simple calculation; |
| | 7 | - creation of a need for formulas;  
- step towards the notion of varying quantity: use of a direct relation in software for doing calculations on data from tabulated values; |
| 3 The Vacuum Pump | 9 | - process of numerical integration on a numerical level;  
- notion of varying quantity and operational introduction of variable and constant;  
- (re)construction of simple direct relations;  
- relation between Δ-formula and stock-flow diagram;  
- formal introduction of process of numerical integration;  
- first use of a direct relation in a computer model;  
- second use of a direct relation in a computer model; |
| 4 Sound | 15 | - practising with the concepts of variable and constant;  
- second use of a direct relation in a computer model;  
- introduction of networks of direct relations in graphical models; |
| 5 Force and Movement | 16 | - possibly twofold and threefold roles of a variable; dummy variables;  
- adaptation of Δ-formulas to graphical modelling;  
- construction of a simple Δ-formula for a new realistic situation;  
- integration of all elements required for graphical modelling;  
- constructing a model based on known equations; |
6.6 General setting and research instruments

The designed instructional materials were used in all seven second year classes and all six third year classes of the Montessori Lyceum of the Hague (HML). Students participated in senior general secondary education or pre-university education in mixed classes. Approximately 40% of all Dutch students of this age participate in these two levels of secondary education. The focus for this paper is on the third year (age: 14-15 years). In two of the six third year classes, the first author of this paper was also the teacher. A third class was taught by an experienced teacher who had also participated as a teacher at an earlier stage of the project. The other three classes were taught by three different teachers who were all novices, both to teaching physics and to modelling. From two of these teachers, we often did not receive data.

In general, two types of research instruments have been used:
1. in order to study the effects of the instructional materials, to study students’ conceptions and changes of these conceptions, to detect student difficulties, and in order to be able to relate test results to student behaviour, we made classroom observations, audio recordings, and screen recordings of students working with the instructional materials, especially the computer tasks; in addition, the results of the computer tasks have been collected. Students worked individually or in small groups with the instructional materials.
2. in order to evaluate students’ understanding of the content of each module, final tests has been given to the students. Part of the questions in these tests are designed to address the research questions.

The research instruments are described more specifically, together with the outcomes of field testing, in Section 6.7 for the module Resistance and Conductance, in Section 6.8 for the module The Vacuum Pump, and in Section 6.9 for the module Force and Movement. The module Sound is not discussed in this paper; results from field testing of this module will only occasionally be referred to. An overview of the research instruments in relation to the research questions listed in Section 6.3 is shown in Table 6.2.

Education at HML is based on principles of Montessori education. Some special features of this education must be accounted for when analysing and interpreting test results. To a certain extent, students are allowed to work at their own pace. This puts extra demands on the instructional materials: students must be able to work with these materials without much help of the teacher. Therefore they usually are provided with materials for checking their own work (but not the final tests). This makes written assignments comparatively less useful for research purposes. There are more discussions among students and between the teacher and small student groups than in traditional teaching. Relatively much research information comes from observations and audio recordings of these discussions.
Often (but not always) students are allowed to postpone a test until they are ready for it. Also, students are allowed to redo a test if they failed the first time. Thus, more than in traditional teaching, tests have a formative character and several versions of the final test are needed for each module.

Sometimes, students finish parts of a module only after doing a test, although this order usually is dissuaded by the teachers. Therefore, test results are not always final results of instruction and learning. In our analysis, we must carefully check whether important student work has been carried out before a test is done. In some cases, we must decide which test attempt of a student to use for analysis. Generally, we use test results of the first attempt, for pragmatic reasons and in order to avoid a bias towards good results.

Near the end of the third year of secondary education, students in the Netherlands must have chosen a set of courses for the next year. Sciences need not to be part of this set. Motivation of students not applying for sciences often declines. At traditional schools, this may lead to bad test results for these less motivated students at the end of the third year. At HML, it often leads to a decreasing number of students participating in the tests towards the end of the year.

### Table 6.2

<table>
<thead>
<tr>
<th>Module</th>
<th>Research instrument</th>
<th>Research Question</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Computer formula task</td>
<td>1  2  3  5  4  6</td>
</tr>
<tr>
<td>Resistance and Conductance</td>
<td>Classification test question</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Computer task test question</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Construction test question</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pump model computer task</td>
<td><strong>x  x  x  x</strong></td>
</tr>
<tr>
<td></td>
<td>Interest model computer task</td>
<td><strong>x  x  x  x</strong></td>
</tr>
<tr>
<td></td>
<td>Pump model test questions</td>
<td><strong>x  x  x</strong></td>
</tr>
<tr>
<td></td>
<td>Variable classification test question</td>
<td><strong>x</strong></td>
</tr>
<tr>
<td></td>
<td>Diagram to Δ-formula test question</td>
<td><strong>x</strong></td>
</tr>
<tr>
<td></td>
<td>Construction test question 1</td>
<td><strong>x</strong></td>
</tr>
<tr>
<td></td>
<td>Construction test question 2</td>
<td><strong>x</strong></td>
</tr>
<tr>
<td>Force and Movement</td>
<td>Computer modelling tasks</td>
<td><strong>x  x</strong></td>
</tr>
<tr>
<td></td>
<td>Direct relation construction test question</td>
<td><strong>x  x</strong></td>
</tr>
<tr>
<td></td>
<td>Δ-formula construction test question</td>
<td><strong>x  x  x</strong></td>
</tr>
<tr>
<td></td>
<td>Δ-formula to diagram test question</td>
<td><strong>x</strong></td>
</tr>
</tbody>
</table>

Data from the test questions are students’ answers, data from the computer tasks consist of handed-in results, screen recordings, audio recordings, and observation notes about students at work.

---

**Table 6.2**

Often (but not always) students are allowed to postpone a test until they are ready for it. Also, students are allowed to redo a test if they failed the first time. Thus, more than in traditional teaching, tests have a formative character and several versions of the final test are needed for each module.

Sometimes, students finish parts of a module only after doing a test, although this order usually is dissuaded by the teachers. Therefore, test results are not always final results of instruction and learning. In our analysis, we must carefully check whether important student work has been carried out before a test is done. In some cases, we must decide which test attempt of a student to use for analysis. Generally, we use test results of the first attempt, for pragmatic reasons and in order to avoid a bias towards good results.

Near the end of the third year of secondary education, students in the Netherlands must have chosen a set of courses for the next year. Sciences need not to be part of this set. Motivation of students not applying for sciences often declines. At traditional schools, this may lead to bad test results for these less motivated students at the end of the third year. At HML, it often leads to a decreasing number of students participating in the tests towards the end of the year.
6.7 Design, research instruments, and results of field testing of *Resistance and Conductance*

### 6.7.1 Instructional design of the module *Resistance and Conductance*

Steps 6 to 8 of the design of the partial learning path on formulas and variables described in Section 6.4.2 are implemented in the module *Resistance and Conductance*. In the preceding grade, the concepts of electrical current, voltage, resistance, and conductance have already been introduced qualitatively to the students. Therefore, a qualitative notion of these concepts is the starting point of the process of mathematization. The question is how to proceed this process with students who probably are not yet sufficiently familiarized with the structural role of mathematics in physical thought, as identified by Pietrocola (2008). Up to this point of the learning path, most formulas have been introduced starting from numerical examples that subsequently have been generalized. Therefore, we continue with a more or less similar numerical approach. Students are offered three options for a (word) formula relating current, voltage, and resistance. They are asked to select one of these options by comparing numerical results of several calculations with qualitative ideas. Only outcomes of calculations with the correct formula ($R = \frac{U}{I}$) match with a qualitative notion of resistance. Subsequently, the formula for the conductance $G$ is introduced as the inverse formula ($G = \frac{I}{U} = \frac{1}{R}$). Hereafter, students are provided with the operational definitions of formula, equation, and simple calculation, and practise distinguishing between these three concepts (Step 6).

The core task of the module starts with an experiment in which students measure the electrical current as a function of voltage for an electrical device. Results are tabulated in the computer learning environment. Students let the computer calculate the resistance as a function of voltage by entering the formula for resistance. To do so, copying the steps shown in an instructional video suffices. The final step is to let the software calculate the conductance as a function of voltage. For this, the instructional video cannot be copied exactly: students must have grasped the idea of using a formula if they are to accomplish this last task (Steps 7 and 8). This task is called the computer formula task.

We expect the students to be fully occupied with understanding the processes of measuring, managing the software, and understanding the use of formulas in the software. The understanding of the physics content, i.e., the graphs and their interpretation, is a learning goal of a subsequent section of the module.

### 6.7.2 Description of the test questions

There are two versions of the final test of the module. Three questions in each test are used for research purposes.
6.7.2.1 Classification test question. This test question consists of four expressions to be classified as formula, equation, or simple calculation, according to the operational definitions given in Step 6 of Section 6.4.2. All expressions are deliberately based on formulas that the students have used the year before.

6.7.2.2 Computer task test question. The test question resembles part of the computer formula task. It is used for measuring students’ skills and understanding with respect to the use of formulas in the computer learning environment. Students are given an already filled-in table and a screen capture of a dialog window similar to the ones they have used in the computer formula task (see Figure 6.2). A beginning of the calculation for the first row of data is given at a numerical level, so students may generalize this calculation into a formula. Students are asked what must be entered in the dialog window. To investigate whether transfer has taken place, the data in the table and the formula are not chosen from electricity, but from fields of physics that students have studied the year before. In order to investigate if the students are not merely memorizing the formula, in one of the versions of the test, unusual symbols are used in the header of the table (L and T instead of Δx and Δt). In their answers to the test question, students are expected to use these unusual symbols too.

![Figure 6.2](replica.png)

Figure 6.2: Screen capture of the computer formula task (replica in English of Dutch original).

6.7.2.3 Construction test question. As mentioned in Section 6.7.1, a numerical approach is used for the introduction of the formula relating resistance to voltage and current. In this approach, the correct (word) formula must be chosen by comparing qualitative ideas with numerical outcomes from calculations. The construction test question is meant to explore to what extent these students, who never have constructed a formula without assistance of a teacher before, have adopted this approach. Students must choose a formula for heat conductance of a wall (test version 1) or for flow resistance for water through a tap (test version 2) from multiple options. In the text of the question, relations between the relevant quantities are described qualitatively. Students are asked to explain their choices.
6.7.3 Outcomes of the computer formula task of the module Resistance and Conductance

6.7.3.1 Handed-in results of the computer formula task. We collected 45 results of the computer formula task of 99 out of 138 students (72%). A number of students had not finished this task completely, i.e., including the graph for \( G \). But in the 29 results that are complete, all formulas are correct. Only two errors with respect to the name of the quantity ('Formula 1' instead of 'G') and two errors with respect to the unit occur.

6.7.3.2 Audio recordings and observations of the computer formula task. As expected, students focussed on the use of the formula in the computer learning environment and not on the physics content of the task. They seemed to handle the computer formula task quite smoothly, but we also observed and recorded some student problems. Most of these problems had to do with a lack of basic knowledge with respect to the concepts, the symbols, the distinction between quantities and units, and the required formula. Some students had not finished or understood the module’s preceding sections, in which the required formulas are introduced. A few students tried to enter \( U/I \) for \( G \), instead of \( I/U \); probably they took the instructional video too literally. This suggests that these students did not grasp the required mathematical and physical concepts yet.

6.7.3.3 Screen recordings of the computer formula task. We made 7 screen recordings involving 15 students working on the computer formula task. In all recordings, student difficulties with respect to the dialog window occur, in spite of the attention for the format of this window in the instructional video. An example is shown in Figure 6.2. Students enter a formula conform the operational definition in the instructional materials into the ‘Formula’ field, whereas the left part of this formula should be entered in the ‘Quantity’ field and the equal-sign must be left out. This is an example of a mismatch between mathematical notation and input notation of a computer learning environment. Some student errors suggest that not all students are able to distinguish between a physical quantity and a unit.

6.7.4 Outcomes of the test questions of the module Resistance and Conductance

In order to compare students’ performances on the test questions with their performances on the computer formula task, five categories, labelled A to E, of students are distinguished with respect to this task. These categories are described in the first column of Table 6.4, together with the distribution of students over these categories in the second column. Categories A and B need clarification. Category A is the least informative because students in this category did not hand in the computer formula task. Therefore we do not know
if or how they carried out this task. Students in Category B handed in the result of their computer formula task only some weeks after the test. They all belong to one class in which many students did the test in spite of not being ready for it. From these students, we can be almost sure that they did not carry out the computer formula task before the test. Part of them redid the test later.

6.7.4.1 Outcomes of the classification test question. We received 94 student results from four classes distributed over two versions of the final test of Resistance and Conductance. The distributions of students’ classifications are listed in Table 6.3. Formulas only consisting of symbols are classified correctly by three quarters of all students. Equations and simple calculations are rarely classified as formulas (approximately 10%). The expressions that are incorrectly classified most often are formulas containing two symbols and one number. Probably, part of the students have been troubled by the alternative conception, expressed by some of these students on another occasion, that a formula must only consist of symbols. There is also some confusion between equations and simple calculations (approximately 20%). We conclude that students are able to grasp the notion of formula according to our operational definition, although students’ base of orientation still needs some improvement.

<table>
<thead>
<tr>
<th>Expression</th>
<th>simple calculation</th>
<th>equation</th>
<th>formula</th>
<th>no answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta x = 27.8 \cdot 5.2$</td>
<td>60%</td>
<td>25%</td>
<td>12%</td>
<td>4%</td>
</tr>
<tr>
<td>$12 = \rho \cdot 3.2$</td>
<td>21%</td>
<td>60%</td>
<td>8%</td>
<td>12%</td>
</tr>
<tr>
<td>$P = \Delta E / \Delta t$</td>
<td>15%</td>
<td>9%</td>
<td>74%</td>
<td>2%</td>
</tr>
<tr>
<td>$F_z = m \cdot 9.8$</td>
<td>6%</td>
<td>50%</td>
<td>35%</td>
<td>10%</td>
</tr>
<tr>
<td>Version 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho = 23 / 7$</td>
<td>76%</td>
<td>12%</td>
<td>10%</td>
<td>2%</td>
</tr>
<tr>
<td>$120 = v_{av} \cdot 8.7$</td>
<td>26%</td>
<td>52%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>$\Delta E = P \cdot \Delta t$</td>
<td>2%</td>
<td>19%</td>
<td>79%</td>
<td>0%</td>
</tr>
<tr>
<td>$m = F_z / 9.8$</td>
<td>19%</td>
<td>40%</td>
<td>21%</td>
<td>20%</td>
</tr>
</tbody>
</table>

Distribution of students’ classifications for the classification question of test versions 1 (52 students) and 2 (42 students) of the module Resistance and Conductance. In the first column the expressions are listed. How students classified these expressions is shown in the other columns. Correct classifications are shaded.
In Table 6.4, students’ performance on this classification test question is compared with their category with respect to the computer formula task. On average, students from Category E perform better on this classification test question than students from other categories. This suggests that participation in the computer formula task positively contributes to students’ conception of formula. That this participation must be active follows from the average result of students from Category D, which is below average. Just imitating the instructional video is not sufficient.

### Table 6.4

**Result of the computer formula task vs. quality of students’ classifications of expressions**

<table>
<thead>
<tr>
<th>Student category</th>
<th>Number of students</th>
<th>Number of correct classifications</th>
<th>Average score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A. No result handed-in</td>
<td>30 (32%)</td>
<td>1 (1%)</td>
<td>5 (5%)</td>
</tr>
<tr>
<td>B. Result handed-in after the final test</td>
<td>11 (12%)</td>
<td>1 (1%)</td>
<td>3 (3%)</td>
</tr>
<tr>
<td>C. Incomplete result in which no formula has been used</td>
<td>8 (9%)</td>
<td>2 (2%)</td>
<td>2 (2%)</td>
</tr>
<tr>
<td>D. Only the formula is used for which copying the instructional video suffices</td>
<td>18 (19%)</td>
<td>5 (5%)</td>
<td>8 (9%)</td>
</tr>
<tr>
<td>E. Computer formula task completed</td>
<td>27 (29%)</td>
<td>8 (9%)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>All categories</td>
<td>94 (100%)</td>
<td>2 (2%)</td>
<td>23 (24%)</td>
</tr>
</tbody>
</table>

In the first column, the student categories with respect to the computer formula task are listed; the numbers of students per category are shown in the second column. In the next five columns, the number of students per category and per number of correct classifications in the classification test question is presented. In the last column the average score on the classification test question for each student category is shown.

In Table 6.4, students’ performance on this classification test question is compared with their category with respect to the computer formula task. On average, students from Category E perform better on this classification test question than students from other categories. This suggests that participation in the computer formula task positively contributes to students’ conception of formula. That this participation must be active follows from the average result of students from Category D, which is below average. Just imitating the instructional video is not sufficient.

**6.7.4.2 Outcomes of the computer task test question.** Students’ results on the question what should be entered into the ‘Formula’ input field of the dialog window are presented in Table 6.5, both for all 94 students and per category of the computer formula task. Almost half of all students give an answer that can be considered as correct or ‘having the correct intention’. Fully correct are answers that form the right part of the formula that must be
entered. Only 10 students gave such an answer. An answer consisting of the complete formula, including a left part and an equal sign (cf. Figure 6.2) has been given by 19 students. For these students, the operational definition prevails above the notation for the computer task. Other answers that are considered to have the correct intention are word formulas, formulas expressed in units instead of symbols for variables, formulas in which upper case and lower case letters are confused, and formulas consisting of other symbols than the symbols used in the table from the test question. If these incorrect formats are used in the real learning environment, the program will not run.

In spite of the operational definition of formula presented to them, 10% of all students use a number or an expression only consisting of numbers instead of a formula.

The percentage of correct intended answers in Category E is greater than that for all students together, and remarkably greater than the corresponding percentage in Category D. This suggests that actively participating in the complete computer formula task is important for the acquisition of the ability to use a formula in the computer learning environment. Just imitating the instructional video is not sufficient.

### 6.7.4.3 Outcomes of the construction test question

Results of the construction test question are presented in Table 6.6. Students’ explanations vary considerably. Many students give no explanation at all, give an unclear explanation, or merely quote parts of the text of the question. Most correct answers are based on the analogy detected by students between the test question and the similar case from electricity. Even for some incorrect answers, the analogy is correct, but an incorrect formula from electricity is mimicked. Only a few students use more sophisticated arguments, applying calculations combined with qualitative proportionality arguments effectively. Apparently, most

<table>
<thead>
<tr>
<th>Answer to the computer task test question</th>
<th>Student category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>'Correct intention'</td>
<td>11 (37%)</td>
</tr>
<tr>
<td>Incorrect</td>
<td>8 (27%)</td>
</tr>
<tr>
<td>No answer</td>
<td>11 (37%)</td>
</tr>
<tr>
<td>Total</td>
<td>30 (100%)</td>
</tr>
</tbody>
</table>

Student categories A to E are described in Table 6.4.
students have not yet adopted the construction approach from the instructional materials. If we want our students to learn to use such approaches themselves, specific training is required. Analogy based reasoning seems to be a good candidate as an alternative method for the construction of formulas, but this can be questioned. Some students clearly base their answer on a linguistic analogy and the ill-prepared students from the class mentioned at the beginning of Section 6.7.4 do not perform worse on this question. Students with a poorer mathematical and physical understanding probably successfully rely on analogical reasoning, while the mathematical understanding of the other students is not yet good enough for these abstract formulas at this stage of their development.

### Summary of the outcomes of Resistance and Conductance

Most students at the beginning of the third year of secondary education are able to classify formulas as such using the operational definition given to them, although the students’ base of orientation still needs some improvement (Section 6.7.4.1). Results of the computer formula task and the computer task test question confirm the idea that using a formula in the computer learning environment correctly is not easy for students at this stage of their development (Sections 6.7.3 and 6.7.4.2). Results also suggest that actively participating in the computer activity positively contributes to both the ability of using formulas (Section 6.7.4.2) and an understanding of the notion of formula (Section 6.7.4.1). Instrumental issues may have negatively influenced test results.

<table>
<thead>
<tr>
<th>Test version and formula</th>
<th>Student answer</th>
<th>Number of student answers</th>
<th>Analogy based answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 1, on conductance for heat of a wall</td>
<td>correct</td>
<td>23 (44%)</td>
<td>4</td>
</tr>
<tr>
<td>incorrect</td>
<td>28 (54%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no answer</td>
<td>1 (2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>52 (100%)</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Version 2, on flow resistance of a tap</td>
<td>correct</td>
<td>32 (76%)</td>
<td>19</td>
</tr>
<tr>
<td>incorrect</td>
<td>9 (21%)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>no answer</td>
<td>1 (2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>42 (100%)</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>

The last column shows the number of student answers that is based on the analogy with the corresponding formula from electricity.

Table 6.6

6.7.5 Summary of the outcomes of Resistance and Conductance

Most students at the beginning of the third year of secondary education are able to classify formulas as such using the operational definition given to them, although the students’ base of orientation still needs some improvement (Section 6.7.4.1). Results of the computer formula task and the computer task test question confirm the idea that using a formula in the computer learning environment correctly is not easy for students at this stage of their development (Sections 6.7.3 and 6.7.4.2). Results also suggest that actively participating in the computer activity positively contributes to both the ability of using formulas (Section 6.7.4.2) and an understanding of the notion of formula (Section 6.7.4.1). Instrumental issues may have negatively influenced test results.
Outcomes from both the construction test question and the computer task test question indicate that students are in need of strategies for the construction of formulas. Results on these test questions can be considered promising if it is realised that these students have never constructed formulas without teacher assistance before.

6.8 Design and results of field testing of the module The Vacuum Pump

6.8.1 Instructional design of the module The Vacuum Pump

Steps 9 to 14 of the design described in Section 6.4.2 are implemented in the module The Vacuum Pump. This module is mainly the same as the previous version described by Van Buuren et al. (2012). After performing measurements with a manually driven vacuum pump, students must complete a graphical model for this pump by entering the formula for \( N_p \), the number of molecules to be pumped out per beat, as a function of \( N_{\text{tot}} \), the total number of molecules in the pump and vessel together. In order to do so, students must also understand how the pump factor (the fraction of molecules pumped out each beat) depends on the volumes of pump and vessel. This computer modelling task is referred to as the pump model computer task. The main differences between this version and the previous version of this module are the following:

1. The concepts of variable and constant are introduced explicitly before the pump model computer task (Step 9).
2. In order to reduce cognitive load during the pump model computer task, introductory and more general student activities have been transferred from the actual computer task to the written materials as much as possible, to be studied before, and in some cases after the computer task.
3. Students are provided with a few extra exercises on the determination of the pump factor before the start of the pump model computer task. This is expected to reduce the cognitive load for the phase in which the formula for \( N_p \) must be reconstructed, but not the load of the task of using a formula in the computer learning environment.
4. Two exercises are added to practise the construction of direct relations (see Step 10).
5. The relation between the stock-flow diagrams and \( \Delta \)-formulas is introduced (Step 11). This relation has also been shown to the students at the end of the preceding grade, but many students did not advance far enough that year or did not grasp the idea. Two exercises on the translation of stock-flow diagrams into \( \Delta \)-formulas and vice versa are added to the instructional materials.
6. The pump model computer task is extended with an exercise in which the vacuum pump used in the numerical introductory example in the instructional
materials must be modelled before the real pump is modelled. Recognition of the numbers from the computer output of this model is expected to enhance students’ confidence in the model and their own modelling skills.

7. The ‘interest model computer task’ is added (Step 14). Students must build a model for a bank-account (Figure 6.3). Each month, a constant amount of money $M$ is deposited into this account, but the capital $C$ on this account also grows because of a monthly interest $I$. Students must construct the direct relation for the interest.

![Figure 6.3: Model for a bank account. The capital $C$ grows as result of a monthly deposit $M$ and interest $I$. Students must construct the direct relation for $I$.](image)

6.8.2 Description of the test questions

There are three different but similar versions of the final test of this module. Several questions in this test are used for research purposes.

**6.8.2.1 Pump model test questions.** In order to be able to compare the current version of the learning path with the previous version, the pump model test questions are the same as the set of questions used by Van Buuren et al. (2012). In the main pump model test question, students are asked to write down the formula for the calculation of the number of molecules $N_p$ in the pump (see Section 6.8.1). Before this main question, students are asked to manually calculate the number of molecules in the pump and the vacuum vessel for a few pump beats. For this, they need to determine the pump factor. The three versions of the test differ with respect to the initial value for $N_{tot}$ and the dimensions of the pump and the vessel.

**6.8.2.2 Variable classification test question.** The variable classification test question is meant to measure the ability to distinguish between variables and constants. Students are given experimental results and model output for an unknown vacuum pump. The output consists of graphs of $N_{tot}$ as a function of the number of pumping beats. Students are asked to discuss these graphs. Two sub questions are about two of the three relevant quantities ($N_{tot}$, $N_p$, and the pump factor). Students are asked whether these are variables or constants. Differences between the three test versions are in the selection of two of these three quantities, but the pump factor was present in all test versions.
6.8.2.3 **Diagram to \(\Delta\)-formula test question.** The diagram to \(\Delta\)-formula questions measures the ability to interpret a stock-flow diagram as a \(\Delta\)-formula. Students are asked to give the formula corresponding to the stock-flow diagram shown in Figure 6.4. This question is part of an overall question in which students’ mathematical understanding of graphical models is probed.

![Diagram to \(\Delta\)-formula question in the final test of The Vacuum Pump](image)

Figure 6.4: Figure from the diagram to \(\Delta\)-formula question in the final test of *The Vacuum Pump*.

6.8.2.4 **Construction test questions 1 and 2.** A consequence of our new approach may be that, if students’ results of the computer task and the pump questions have improved, this may just be the effect of more practising and/or memorisation. In order to investigate to what extent transfer has taken place, construction questions 1 and 2 are added to the test. These two questions measure the ability to construct direct relations for new situations. Both questions are part of the overall question mentioned in Section 6.8.2.3. Students are given a simplified and incomplete model for the increase of the amount of mice (or, in the other test versions, rabbits and hamsters) per month (Figure 6.5).

In construction question 1 students are asked what must be entered in this model for the amount \(Y\) of new-born mice per month. This direct relation, the product of the total amount of mice \(M\) and the birth rate, must be constructed by the students. This construction resembles to some extent the (re)construction part of the computer task from *The Vacuum Pump*.

![Construction question 1 in the final test of The Vacuum Pump](image)

Figure 6.5: Figures from construction question 1 in the final test of *The Vacuum Pump*. Students are asked what must be entered next to ‘\(Y\) =’. This should be the product of the total amount of mice \(M\) and the birth rate.

In construction question 2, students are asked to construct a formula that gives the relation between the total amount of money \(A\) that a breeder gets...
Learning to use and create formulas for constructing computer models

when he sells N mice for a price p per mouse. This question resembles one of the two exercises in the instructional materials (Section 6.8.1, Point 4).

Beforehand, we expect construction test question 1 to be more difficult than construction test question 2, because question 1 is closely related to the graphical model, in which the Δ-formula from the diagram to Δ-formula test question (Section 6.8.2.3) plays a part too. This may confuse students.

6.8.3 Outcomes of the learning tasks from the module The Vacuum Pump

6.8.3.1 Observations and recordings of learning tasks preparing for the pump model computer task. Understanding the calculation process on a numerical level was not problematic for the students. This agrees with the findings of Van Buuren et al. (2012). We observed a few students making a remarkable mistake. Although these students understood that at each pump beat the same fraction of molecules is pumped out and although they applied this correctly to the first pump beat, they did not adopt the same approach for the next beats. Being asked to explain their approach, they could not understand why they made this error. Apparently, understanding the ideas is not enough for internalization for these students at that moment.

6.8.3.2 Handed-in results of the pump model computer task. We received 50 results of the computer task from 90 students. In 94% of these results (96% of the students), the formula for \( N_p \) can be considered as correct. This percentage is greater than the percentage of 81% (79% of the students) reported for the previous version (Van Buuren et al., 2012). The error to define \( N_p \) only by means of numbers or expressions consisting of numbers has been made by only two students. The pump factor is correctly adapted to fit the experimental data from the real vacuum pump in 72% of all handed-in results; in 16% of all results, the pump factor has been calculated correctly but has not been adapted for the model to fit the experimental data.

6.8.3.3 Observations and recordings of the pump model computer task. Most students collaborated on the pump model computer task, often explaining things to each other. Students regularly needed to fall back on the preceding sections of the module. Those students who had severe problems usually had not finished these preceding sections. This shows the importance of limitation of the number of learning goals for a modelling activity and of the introduction of important concepts before the actual computer task is carried out. The same holds for the introduction of new features of the software. Some students said that they did not feel confident with the software yet. An example is the use of * and / symbols for multiplication and division, although the number of students having problems with this notation was clearly decreased as a result of the earlier introduction of these symbols.
In spite of the attention paid to the concepts of formula and variable, observations revealed that many students initially still did not use a formula, but entered a simple calculation or, more often, the outcome of this calculation, in the dialog window. In a class consisting of 30 students we made audio recordings of 4 such cases, involving 11 students.

As mentioned in Section 6.2, it can mislead students that for the stock variable no formula must be entered, but only an (initial) value. This is illustrated by a remark of a student, just having understood that for the flow $N_p$ a formula is required: “But then the other one [the stock] must also be wrong, because each beat there are not 500 leaving this thing, or something like that?”

Yet, the concepts of formula, simple calculation, variable, and constant proved to be useful for the learning process. A typical example is the following discussion between a student (S) and the teacher (T). The student has entered $0.1 \times 500$ as definition for $N_p$ instead of $0.1 \times \dot{N}_{tot}$.

T: “If you enter $0.1 \times 500$, do you have a formula, an equation, or a simple calculation?”
S: “A simple calculation, I think.”
T: “A simple calculation. Is its outcome a variable or a constant?”
S: “A constant”
T: “Okay”
S: “and it must not be a constant.”
T: “It must not be a constant.”
S: “But, I am thinking, how can this be done?”

What follows is a discussion on the use of symbols in order to make the expression more general. In this situation, the notions of simple calculation, constant, and variable helped to create a conceptual conflict and a need for a formula consisting of variables represented by symbols.

Modelling requires a shift in focus from outcomes to methods. This is illustrated by the reaction of a student, after being explained that in a computer model it may be more efficient to enter an expression of numbers instead of its outcome: “But a result must always be reduced, isn’t it? It must always be made as simple as possible?!” This focus on outcomes instead of methods can be an obstacle for the learning of modelling. On the other hand, the example also shows that modelling may help to change this focus.

As mentioned in Section 6.8.1, Point 6, an exercise has been added to the computer task in which the vacuum pump from the numerical introductory example in the instructional materials must be modelled. From a few student interviews, the expectation that this addition would enhance students’ confidence in the model and their own modelling skills proved to be correct. Finally, in an interview after the pump model computer task some of the students explicitly mentioned that, during the computer task, they felt a tendency to stop thinking and start guessing.
6.8.3.4 Observations and recordings of the interest model computer task. Only a minority of the students finished the additional computer task (on interest). Part of these students initially entered a number instead of a formula in the dialog window of a variable, as we observed in classroom and on the three screen recordings (involving 6 students), and as can be heard on audio recordings. In 4 out of 9 handed-in results (7 out of 18 students), this variable still is defined by a number only. Two of these four results, however, show that students have correctly answered an earlier question, in which they are explicitly asked to construct this formula. More students may have done so; not all student results contain the answers to earlier questions. Apparently, not being able to construct the formula is not the only reason for the use of numbers instead of formulas in a graphical model.

6.8.4 Outcomes of the test questions from the module The Vacuum Pump

6.8.4.1 Outcomes of the pump model test questions. In Table 6.7 the results of the pump model test questions are presented, both for the previous version (Van Buuren et al., 2012) and for the current version of the module. For the current version, in this table both the results for all 109 students and the results for those 70 students from whom we have also received a result of the computer task are presented.

There are almost no differences between the test results of the two versions for the questions that can be dealt with on a numerical level: the determination of the pump factor and the manual calculated values for the pumping process. There only is a decrease in the number of ‘reasonable’ calculation errors in favour of the number of fully correct answers. This may be due to extra exercises in the instructional materials, but it can also be the result of the recently increased general attention in Dutch secondary education for arithmetic.

With respect to the main question, in which students are asked to give the formula for \( N_p \), the percentage of correct answers has more than doubled. In addition, the quality of incorrect answers is better. Only 3 students have given an answer consisting of numbers instead of a formula. Finally, for the students of whom we can be more sure that they finished the computer task, the percentage of correct answers is somewhat greater than for the entire group of students. Although the relative number of students being able to reconstruct this formula clearly is greater than it was for the previous version of the module, it is still less than 50%. Apparently, our approach has contributed to a better understanding, but reconstruction of this formula is still difficult for many students.
### Table 6.7

**Test results for the pump test questions**

<table>
<thead>
<tr>
<th>Determination of pump factor</th>
<th>Manual calculated values for pumping process</th>
<th>Formula for ( N_p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test results from 80 students from whom the result of the computer task was received, from the previous version of the learning path (Van Buuren et al., 2012).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct or consequent</td>
<td>Correct 49% 59% 16%</td>
<td>‘Reasonable’ calculation error 23%</td>
</tr>
<tr>
<td>Incorrect 19%*</td>
<td>29% 44%</td>
<td></td>
</tr>
<tr>
<td>No answer</td>
<td>10% 13% 40%</td>
<td></td>
</tr>
<tr>
<td>*13% used the pump factor from the example in the worksheet.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test results from all 109 students (current version of the learning path).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct or consequent</td>
<td>Correct 62% 63% 36%</td>
<td>‘Reasonable’ calculation error 10%</td>
</tr>
<tr>
<td>Pump factor from the example in the worksheet is used</td>
<td>18% 18% 6%</td>
<td></td>
</tr>
<tr>
<td>Incorrect</td>
<td>7% 14% 43%</td>
<td></td>
</tr>
<tr>
<td>No answer</td>
<td>2% 5% 16%</td>
<td></td>
</tr>
<tr>
<td>Test results from 70 students from whom the result of the computer task has been received (current version of the learning path).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct or consequent</td>
<td>Correct 63% 66% 40%</td>
<td>‘Reasonable’ calculation error 10%</td>
</tr>
<tr>
<td>Pump factor from the example in the worksheet is used</td>
<td>17% 17% 9%</td>
<td></td>
</tr>
<tr>
<td>Incorrect</td>
<td>7% 11% 41%</td>
<td></td>
</tr>
<tr>
<td>No answer</td>
<td>3% 6% 10%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6.8

**Results of the test questions on the distinction between variables and constants**

<table>
<thead>
<tr>
<th>Is … a variable or a constant?</th>
<th>( N_{tot} )</th>
<th>( N_p )</th>
<th>pump factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of collected student answers</td>
<td>46</td>
<td>63</td>
<td>109</td>
</tr>
<tr>
<td>Number of correct answers</td>
<td>40</td>
<td>48</td>
<td>82</td>
</tr>
<tr>
<td>Percentage of correct answers</td>
<td>87%</td>
<td>76%</td>
<td>75%</td>
</tr>
</tbody>
</table>

Each student was asked to classify two of these quantities.
6.8.4.2 Outcomes of the variable classification test question. Results from the test question in which students must distinguish whether a quantity is a constant or a variable are tabulated in Table 6.8. Each quantity is correctly classified by more than three quarters of all students. Of all 109 students, 74 students (68%) classified both quantities correctly. The total number of molecules \(N_{\text{tot}}\) is more often correctly classified than the pump rate \(N_p\) and the pump factor. This is not surprising. Students may erroneously think that \(N_p\) is a constant, but even in such a case the value of \(N_{\text{tot}}\) must be changing. Errors with respect to the pump factor may be due to the fact that students worked with several different vacuum pumps in the instructional materials. From one pump to another, the pump factor changed. This may have lead some students to the conclusion that the pump factor is a variable. Finally, students may mix up the variable pump rate \(N_p\) and the constant pump factor.

These results do not yet prove that students are able to distinguish between constants and variables in general. The varying or constant character of the quantities in this question has already been discussed in the preparatory phase, and in the test students must choose between only two options, so the questions are not very difficult and positive results can be expected. In classroom we noticed that students had some difficulties with a question about the classification of the electrical resistances \(R\) of a resistor and of a light bulb as variable or constant, but these problems probably were caused by their knowledge of the electrical devices, not by their notion of variable or constant as such. Discussions between teachers and students working on the pump model computer task (Section 6.8.3.3) indicate that students are able to use notions of constant and variable effectively. Therefore, in summary, we conclude that a majority of third year lower secondary students can distinguish between constants and variables.

6.8.4.3 Outcomes of the diagram to \(\Delta\)-formula test question. Answers of all students on the test question in which they are asked to give the formula corresponding to a given stock-flow diagram (shown in Figure 6.4), about the increase of mice (hamsters, rabbits) through birth, are listed in Table 6.9. Only 29% of all students are able to give the correct \(\Delta\)-formula. Closer analysis reveals that students who have handed-in a result on the pump model computer task do not perform better. The most frequent error is the omission of any reference to the time \(t\) or to the time step \(\Delta t\). This may be partially caused by the replacement in the pump model of time steps by pump beats. Many students present a \(\Delta\)-formula taken from examples in the instructional materials, originating from other contexts. Probably, these students know that a stock-flow diagram corresponds to a \(\Delta\)-formula, but they do not know how to link the icons to the \(\Delta\)-formula. Seven students present the formula for the flow variable instead of the \(\Delta\)-formula for the stock variable. This is an example of confusion in case of feedback mechanisms. Eight students give an answer that also seems to be inspired by the formula for the flow, but in which the coefficient (the birth
factor) is connected to the wrong quantity. This bears some resemblance to the reversed equation error (Section 6.2), but it may as well be a result of some sort of reasoning analogous to the pump question, where a variable was also defined by means of the product of the other variable and some factor. Almost no students gave an answer consisting of numbers only.

The great diversity in incorrect student answers suggests that the question was not clear for our students at this stage of their development. One cause may be that there is no real need for students to understand the relation between stock-flow diagrams and Δ-formulas at this point of the learning path. For the main part of the module, students are working on the direct relation for the flow variable. A lack of understanding of the stock-flow diagram hardly hampers students’ progress. Apparently, the relation between stock-flow diagrams and Δ-formulas requires more attention in the instruction.

### 6.8.4.4 Outcomes of construction test question 1.

In this question, students are asked to define the flow variable of the stock-flow diagram from Section 6.8.4.3 (Figure 6.5). Table 6.10 shows the results of this question for all students, for the sub group of students who handed in their result of the computer task, and for the sub group of students who correctly answered the similar question for the vacuum pump. Only 20% of all 109 students have answered this question completely correct; a few other answers are almost correct. Both sub groups do not perform much better. It seems that there has

<table>
<thead>
<tr>
<th>Student answers</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completely correct</td>
<td>32</td>
</tr>
<tr>
<td>Correct, apart from a missing Δ-symbol</td>
<td>2</td>
</tr>
<tr>
<td>Formula for another context, from the examples in the instructional materials</td>
<td>17</td>
</tr>
<tr>
<td>Other incorrect answers</td>
<td>43</td>
</tr>
<tr>
<td>No answer</td>
<td>15</td>
</tr>
<tr>
<td>Total number of students</td>
<td>109</td>
</tr>
<tr>
<td>Answers without time or time step</td>
<td>32</td>
</tr>
<tr>
<td>Answers without a Δ-symbol or similar reference to change</td>
<td>12</td>
</tr>
</tbody>
</table>

Students are asked to give the Δ-formula to the stock-flow diagram of Figure 6.4 (or a similar diagram). The correct answer for the mice-version of the test is $\Delta M = \gamma \Delta t$. 

<table>
<thead>
<tr>
<th>Table 6.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.8.4.4 Outcomes of construction test question 1. In this question, students are asked to define the flow variable of the stock-flow diagram from Section 6.8.4.3 (Figure 6.5). Table 6.10 shows the results of this question for all students, for the sub group of students who handed in their result of the computer task, and for the sub group of students who correctly answered the similar question for the vacuum pump. Only 20% of all 109 students have answered this question completely correct; a few other answers are almost correct. Both sub groups do not perform much better. It seems that there has</td>
</tr>
</tbody>
</table>
been little transfer yet. Even if students do recognize the similarity between the diagrams of construction question 1 and the vacuum pump, this does not mean that they see which variables are analogous to each other in the two different contexts, as is shown by the following student answer:

“\[ Y = \cdot M_{\text{tot}} \] In case of a vacuum pump, the pump factor should be at the empty spot, but I don’t know at the moment what to compare the pump factor with.”

Even if students are able to distinguish between constants and variables and understand that a variable must be defined by a formula, this knowledge seems not to be operational yet for all students: 27% of all students defines the flow variable by means of a number only (the birth factor, instead of the birth rate). This is in accordance with the use of numbers by students in the interest model computer task in spite of the correct formula being known to them (see Section 6.8.3.3). One reason may be that students do not realize that the flow in this question is variable and therefore needs a formula as definition. Contrary to the other questions in this test, the term ‘formula’ is not used in this test question. Therefore part of the students may not have been triggered to look for a formula. They may have thought that the flow in this situation is a constant, may have considered this situation as momentary, or may have suffered from the

<table>
<thead>
<tr>
<th>Constructed formula</th>
<th>All students</th>
<th>Sub group I</th>
<th>Sub group II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>22 (20%)</td>
<td>18 (26%)</td>
<td>11 (28%)</td>
</tr>
<tr>
<td>Correct, apart from a minor error</td>
<td>6 (6%)</td>
<td>5 (7%)</td>
<td>3 (8%)</td>
</tr>
<tr>
<td>Number or simple calculation, no variables</td>
<td>29 (27%)</td>
<td>14 (20%)</td>
<td>8 (21%)</td>
</tr>
<tr>
<td>Other incorrect answers</td>
<td>35 (32%)</td>
<td>23 (33%)</td>
<td>12 (21%)</td>
</tr>
<tr>
<td>No answer</td>
<td>17 (16%)</td>
<td>10 (14%)</td>
<td>5 (13%)</td>
</tr>
<tr>
<td>Total number of students</td>
<td>109 (100%)</td>
<td>70 (100%)</td>
<td>39 (100%)</td>
</tr>
<tr>
<td>Answers with some reference to time or time step</td>
<td>9 (8%)</td>
<td>7 (10%)</td>
<td>4 (10%)</td>
</tr>
</tbody>
</table>

Table 6.10

Student answers on construction test question 1, on the construction of a direct relation defining a flow variable

In the first column the characteristics of the constructed formula are listed. In the second column the results of all students are shown. Sub group I consists of students who also handed in their result for the computer formula task. Sub group II consists of students who correctly answered the similar question for the pump model. The formula to be constructed is a direct relation for the increase of animals as a result of birth. The correct answer for the mice-version of the test is \( Y = 2.5 \cdot M \).
misconception that in graphical models numbers are required only. Students’ focus on outcomes may also be a reason for the use of numbers.\textsuperscript{7} In all these cases, the outcome of the multiplication of the birth factor and the initial number of mice would be expected, however. Because an initial number of mice is not given in the text of the test question, students may just have used the only numbers that were available.

Finally, there is a great variety of incorrect answers. Four students give a formula for the stock instead of the flow. Two students convert the formula for the stock into a formula for the flow. Most other answers are hard to interpret, and 17 students did not answer the question at all. This suggests that for many students it is not yet clear how to solve such a task. There are only three indications of the reversed equation error (Section 6.2). Some reference to time has been made by 8\% of all students.

<table>
<thead>
<tr>
<th>Constructed formula</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>66 (61%)</td>
</tr>
<tr>
<td>Almost correct*</td>
<td>6 (6%)</td>
</tr>
<tr>
<td>Incorrect</td>
<td>16 (15%)</td>
</tr>
<tr>
<td>No answer</td>
<td>21 (19%)</td>
</tr>
<tr>
<td>Total</td>
<td>109 (100%)</td>
</tr>
</tbody>
</table>

* apart from the addition of one or two Δ-symbols.

In this question, students must construct the (direct) relation between three variables.

Table 6.11

### Outcomes of construction test question 2

In Table 6.11 the results are listed of all students for construction test question 2, the question on the construction of a formula for the relation between the price $p$ per mouse (rabbit, hamster), the number $N$ of mice sold and the amount $A$ of money received by the breeder. The percentage of correct student answers (61\%) is much greater than the percentage of correct answers to construction question 1. Another 6\% of all students give an answer that is correct apart from the addition of one or two Δ-symbols. Apparently, these students only have problems understanding Δ-notation. All answers contain variables. This is not surprising: no numbers are given, and students have been explicitly asked to

\textsuperscript{7} This may be related to the tendency of students to give one numerical answer to a question and to view an expression not as a proper answer (Booth, 1984, 1988; Chalouh & Herscovics, 1984). This can be considered as a problem with acceptance of lack of closure (Collis, 1975).
Most students present their correct answer as a product \( A = p \cdot N \); a small minority uses a quotient suitable for the calculation of the price \( p \) per animal. Students’ preference for the product version may be explained by the fact that so far in the instructional materials most formulas are presented in product form. The relatively great number of students not giving an answer may be partially explained by the fact that this question is at the end of the test and it is part of a set of four questions. Students may have skipped this question because of lack of time.

6.8.5 Summary of the outcomes of The Vacuum Pump

After having finished The Vacuum Pump, most students can distinguish between a variable and a constant (Section 6.8.4.2). Working on the pump model computer task and the interest model computer task, still many students initially try to define the flow variable by means of numbers only (Sections 6.8.3.3 and 6.8.3.4), but the notions of formula, variable, simple calculation, and constant can be used to change this (Sections 6.8.3.2 and 6.8.3.3). The results of the pump model computer task and the pump model test questions show that students’ use of formulas in modelling has improved compared to the results from Van Buuren et al. (2012). Especially, fewer students use expressions consisting only of numbers and the percentage of students who can reconstruct the formula for the flow in the pump model has more than doubled. Still, more than half of all students is not able to reconstruct this formula, even though many students understand the computational process on a numerical level.

At this stage of their development, approximately 30% of all students can give the \( \Delta \)-formula to a given simple new stock-flow diagram (Section 6.8.4.3). The lack of a clear need to understand the connection between \( \Delta \)-formulas and stock-flow diagrams may be a reason for this. The most important error, made by 30% of all students, is the omission of any reference to the time step \( \Delta t \).

After only a few exercises, almost two thirds of all students are able to construct a simple direct relation connecting three concrete quantities (Section 6.8.4.5), but the interest model computer task (Section 6.8.3.4) and construction test question 1 (Section 6.8.4.4) make clear that there has been little transfer yet with respect to the construction of the direct relation for a flow variable in a graphical model from the pump model to a new but similar situation. Only approximately one-fifth of all students give the required direct relation in construction test question 1. Apparently, at this stage of their development, just a minor part of the students has problems with the construction of simple formulas per se, but a major part of the students has difficulties constructing a direct relation for a flow variable in a graphical model. Although most students are able to distinguish between constants and variables, and understand that a formula must contain variables, a quarter of all students do not yet fully construct a formula, whereas in construction question 1 they have been asked ‘what should be entered’ in the dialog window for the flow variable.
understand that a graphical model is about varying quantities to be defined by formulas.

6.9 Design and results of field testing of the module

*Force and Movement*

6.9.1 Instructional design of the module *Force and Movement*

Steps 16 to 19 of the design described in Section 6.4.2 are implemented in the module *Force and Movement*. For the design of this module, we made use of results of field testing of an earlier version of this module (Van Buuren et al., 2011). In the version of the module that is described in this paper, students combine all elements required for the construction of graphical models for the first time. They get a complete overview of the main elements of graphical models (Step 19). In three computer modelling tasks, students create graphical models starting from given sets of equations. In order to be able to do so, in the instructional materials attention is given to Steps 16 and 17 of Section 6.4.2, that is, on the relation between Δ-formulas and stock-flow diagrams. Students get a few exercises on this relation. Before a new model is implemented in the computer learning environment, students must first design the model on paper. Outcomes of the learning task are discussed in Section 6.9.3, complemented with a few observations from the module *Sound*.

6.9.2 Description of the test questions

There is only one version of the final test of this module. This test can be considered as a final test of the modelling learning path. Three questions are used for the research purposes of this paper.

6.9.2.1 The Δ-formula to diagram test question. In the Δ-formula to diagram test question, a stock-flow diagram must be created for a given Δ-formula. In the instructional materials, students have been practising with such questions. This question is the reverse of the diagram to Δ-formula test question from *The Vacuum Pump* (Section 6.8.2.3).

6.9.2.2 Δ-formula construction test question. The Δ-formula construction test question is part of an overall question in which students’ abilities with respect to the creation of graphical models are investigated. Students are asked to construct a Δ-formula for a situation in which one stock and two flows take part. The stock-flow diagram is not given to the students. The overall question is about forestry: each year, a forester cuts 20% of the trees in a forest and plants a constant number of new trees. Students must construct the Δ-formula for the
Learning to use and create formulas for constructing computer models

number of trees in the forest. The initial number of trees and the constant number of trees that is planted each year are given; students may be tempted to use these numbers. This is the first time that students are asked in a test to construct a Δ-formula for a new situation. In the instructional materials, they have been asked to construct a Δ-formula only once, for a less complicated situation, containing only one flow variable.

6.9.2.3 Direct relation construction test question. The direct relation construction test question is part of the same overall question as the Δ-formula construction test question. Students are asked to construct the formula (a direct relation) for the number of trees that is cut each year. The direct relation is mathematically similar to the relation from construction test question 1 from The Vacuum Pump (Section 6.8.2.4).

6.9.3 Outcomes of the learning tasks of the module Force and Movement

As we observed in classroom and in screen recordings of students working on the computer tasks of Sound and of Force and Movement, formulas have been used in models by most students, although incidentally some students have tried to use numbers only. On an audio recording can be heard how a student, trying to define a variable by means of a number instead of a formula, immediately is corrected by co-workers. In screen recordings of 17 students working on the main computer task of Force and Movement, all but one of the students correctly use formulas. The fact that all models from Force and Movement are based on formulas that are known to the students probably has contributed to this outcome.

6.9.4 Outcomes of the test questions of the module Force and Movement

Because we want to investigate whether students’ abilities with respect to the use of formulas in graphical models have increased, results of these questions are compared to results of the test questions of The Vacuum Pump. Only 27 students took part in the test of Force and Movement, because this module was planned for the end of the third year curriculum and many students did not manage to finish it. Therefore, a bias towards faster and hard-working students can be expected. The 27 students who finished the test were from classes taught by the first author of this paper. Four of these students did the test unprepared, after a break of a few weeks after finishing the work on the module. A fifth student had not carried out the computer tasks before doing the test. From three students we have no test results of The Vacuum Pump. Therefore only 24 student results are used for comparison. In Table 6.12, results of the diagram to Δ-formula test question of The Vacuum Pump are shown for these 24 students. These 24 students indeed performed much better than average on this question (cf. Table 6.9).
6.9.4.1 Outcomes of the Δ-formula to diagram test question. Students’ results on the Δ-formula to diagram test question are shown in Table 6.13. The percentage of correct answers is much greater than in the related (reverse) test question at the end of *The Vacuum Pump* for the same group of students (Table 6.12). In addition, most incorrect answers have been given by ill-prepared students. Therefore, although these 24 students are not completely representative for all third year students, we still conclude that many third year students can create a stock-flow diagram to a given Δ-formula. This does not necessarily mean that students have developed a better understanding of the relation between Δ-formulas and stock-flow diagrams. It may as well be easier to construct a stock-flow diagram for a given formula than, vice versa, to construct a formula to a given diagram.

<table>
<thead>
<tr>
<th>Answer</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completely correct</td>
<td>12 (50%)</td>
</tr>
<tr>
<td>Formula for another context</td>
<td>2 (8%)</td>
</tr>
<tr>
<td>Other incorrect answers</td>
<td>6 (25%)</td>
</tr>
<tr>
<td>No answer</td>
<td>4 (17%)</td>
</tr>
<tr>
<td>Answers without time or time step</td>
<td>3 (13%)</td>
</tr>
<tr>
<td>Answers without a Δ-symbol or similar reference to change</td>
<td>1 (4%)</td>
</tr>
</tbody>
</table>

With respect to the factor time, there is a dissimilarity between stock-flow diagrams and Δ-formulas. In the Δ-formulas, the factor Δt is explicitly present, but in the stock-flow diagrams, this factor is not visualised. In the
Learning to use and create formulas for constructing computer models

6.9.4.2 Outcomes of the Δ-formula construction test question. An overview of students’ answers on the Δ-formula construction test question is shown in Table 6.14. The percentage of completely correct answers on this question is smaller than the percentage of correct answers of the same students on the diagram to Δ-formula test question from The Vacuum Pump (Table 6.12). Yet, we are of opinion that these students have booked some progress. Not only is this Δ-formula construction test question more demanding than the diagram to Δ-formula test question, the quality of incorrect answers is also definitely greater. An example of an almost correct answer is “\[B=50000+(A-K)\cdot \Delta t\]” instead of \[\Delta B=(A-K)\cdot \Delta t\]. Here, \(\Delta B\) has actually been replaced by ‘B minus the initial value’. This answer clearly shows a notion of change of \(B\), but it holds only for the first step of integration. The percentage of students leaving out the factor \(\Delta t\) from the equation is still remarkably great. Even if answers from ill-prepared students are left out, this percentage would just be 25%. Finally, there are indications that the omission of the outflow from the Δ-formula is related to a regularly occurring misconception that an outflow does not influence the stock.

Table 6.14

<table>
<thead>
<tr>
<th>Quality of the constructed Δ-formula</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Δ-formula</td>
<td>9 38%</td>
</tr>
<tr>
<td>‘Outflow’ or ‘Inflow’ missing</td>
<td>3 12%</td>
</tr>
<tr>
<td>Other incorrect answers</td>
<td>12 (5) 50%</td>
</tr>
<tr>
<td>Total</td>
<td>24 (5) 100%</td>
</tr>
<tr>
<td>Formulas without Δt</td>
<td>9 (4) 38%</td>
</tr>
<tr>
<td>Formulas containing an initial value.</td>
<td>7 (4) 29%</td>
</tr>
</tbody>
</table>

The numbers in brackets are the numbers of ill-prepared students.

6.9.4.3 Outcomes of the direct relation construction test question. In Table 6.15 results are presented for the direct relation test question, together with the results of the same students of the similar test question of The Vacuum Pump.
The ability of students to construct a formula seems to have improved: the number of correct answers has more than doubled. The number of answers only consisting of numbers has decreased to almost none, in spite of the presence of more numbers to choose from in the text of the question, and almost all answers contain variables. It must be noticed, however, that in this direct relation construction question students are explicitly asked to construct a formula, whereas in construction test question 1 from *The Vacuum Pump* students were asked ‘what must be entered’ in a dialog window. The percentage of correct answers of the direct relation construction test question is smaller than the percentage of correct answers on construction test question 2 from *The Vacuum Pump*, but the context of that question probably was less difficult for students and it contained less quantities to choose from. Also, the variables in that question could be dealt with on the level of placeholder, whereas two out of three variables in the direct relation construction test question from *Force and Movement* are varying objects. Finally, no indications of the reversed equation error of (Section 6.2) have been found.

### Table 6.15

**Comparison of results from direct relation construction questions from Force and Movement and The Vacuum Pump**

<table>
<thead>
<tr>
<th>Quality of the constructed direct relation</th>
<th>Results from <em>The Vacuum Pump</em></th>
<th>Results from <em>Force and Movement</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Completely correct</td>
<td>5 (21%)</td>
<td>12 (50%)</td>
</tr>
<tr>
<td>Almost correct</td>
<td>1 (4%)</td>
<td>1 (4%)</td>
</tr>
<tr>
<td>Number or simple calculation</td>
<td>6 (25%)</td>
<td>1 (4%)</td>
</tr>
<tr>
<td>Other incorrect answers</td>
<td>11 (46%)</td>
<td>9 (38%)</td>
</tr>
<tr>
<td>No Answer</td>
<td>1 (4%)</td>
<td>1 (4%)</td>
</tr>
<tr>
<td>Total</td>
<td>24 (100%)</td>
<td>24 (100%)</td>
</tr>
<tr>
<td>Answers with reference to time or time step</td>
<td>4 (17%)</td>
<td>2 (8%)</td>
</tr>
</tbody>
</table>

In the second column the results of 24 students on construction test question 1 from *The Vacuum Pump* are shown. The third column contains the results of the same students on the (similar) direct relation construction test question from *Force and Movement*.

### 6.9.5 Summary of the outcomes of Force and Movement

At the end of the third year of secondary education, most students can use direct relations in graphical models (Section 6.9.3). Also, more than two thirds of the students who did the final test of *Force and Movement* are able to construct a stock-flow diagram to a given \( \Delta \)-formula. But the relation between a stock-flow
Learning to use and create formulas for constructing computer models

diagram and time is not clear for an important minority of the students. Almost 30% of the students add an icon for the time step when translating a \( \Delta \)-formula into a stock-flow diagram (Section 6.9.4.1).

Although our group of students may not be representative for all students, results on the construction of a \( \Delta \)-formula for a new situation can be considered promising: almost 40% of the students can construct a \( \Delta \)-formula with limited practising beforehand, the quality of many incorrect answers is fairly good, and most incorrect answers are from ill-prepared students. But here also, there are difficulties with the time step in the formula. Almost two fifth of the students leave out this time step. Student errors also indicate that the role of initial values needs more attention (Section 6.9.4.2). The number of students being able to construct a direct relation for a flow variable in a new situation has increased, but it must be noted that the students were explicitly asked for a formula, contrary to construction test question 1 from *The Vacuum Pump* (Section 6.9.4.3).

6.10 Conclusions, discussion, and recommendations

In Sections 6.10.1 to 6.10.6, all six research questions are answered and results are discussed. In case answers are not satisfactory, recommendations for further research are made. In Section 6.10.7, we reflect on choices we have made for the learning path. In Section 6.10.8, recommendations for further design and research are summarised.

6.10.1 Answer to Research Question 1

*To what extent can students grasp the notion of formula as providing a relation between symbols representing variables and constants?*

Most students at the beginning of the third year of secondary education grasp the notion of formula as consisting of symbolised quantities after explicit instruction: formulas that only consist of symbols are classified correctly by three quarters of all students (Section 6.7.4.1). This notion is sustained or developed further in the learning path. In all other test questions after the introduction of the notions of formula and variable, when asked for a formula, hardly any students have used expressions consisting of numbers anymore (Sections 6.8.4.3, 6.9.4.2, and 6.9.4.3). Yet, students can have difficulty classifying formulas in which both symbolised quantities and a number appear. They can be troubled by the alternative conception that a formula must *only* consist of symbols. The students’ base of orientation needs improvement in order to deal with such subtleties. The results suggest that participating in a computer activity in which formulas must be used to define quantities, contributes both to the ability of using formulas for defining a variable in a
computer environment and to the notion of formula. This participation must be active; merely imitating an instructional video does not suffice (Sections 6.7.4.1 and 6.7.4.2).

6.10.2 Answer to Research Question 2

To what extent can students understand the difference between constants and variables?

Most students can distinguish between variable and constant after the introduction of these concepts in *The Vacuum Pump*. In Section 6.8.4.2, two thirds of all students correctly classified two different quantities as variable or constant. Other support for this conclusion comes from classroom observations (Sections 6.8.3.3 and 6.8.4.2).

6.10.3 Answer to Research Question 3

To what extent can students utilise conceptions of variable and formula for modelling?

The results of the computer formula task and the computer task test question of *Resistance and Conductivity* (Sections 6.7.3 and 7.4.2) are in accordance with the findings of Van Buuren et al. (2012) that the use of formulas containing symbolised quantities in educational software is not self-evident; students need clear notions of formula and variable for modelling. Isolated notions of formula and variable are insufficient: students do not automatically apply these notions when defining a quantity in a computer model (Sections 6.8.3.3, 6.8.3.4, and 6.8.4.4). But clearer notions of formula and variable can be used to change this. The notions of simple calculation and constant, and of formula and variable must be coupled. Students can achieve this with some teacher assistance (Section 6.8.3.3). The following sequence can be effective in a discussion between teacher and student:

1. a quantity defined by a simple calculation actually is a constant,
2. but the quantity should be variable,
3. so we need a formula to define this quantity.

In this way, these clearer notions of formula and variable contribute to students’ understanding and modelling performance: the percentage of correct formulas as an answer to the pump model test question has more than doubled compared to results of Van Buuren et al. (2012), and both the percentage of answers only consisting of numbers and the percentage of students not answering the question have dropped considerably (Section 6.8.4.1).

These clearer and coupled notions of variable and formula are necessary, but at this stage they are not yet sufficient. In the first modules after the formal introduction of formula and variable, more than half of all students still does not
use a formula for defining a flow variable in a graphical model when a formula is not explicitly asked for (Sections 6.8.3.3, 6.8.3.4, and 6.8.4.4). Possible reasons are:
• because the coupling of the notions of formula and variable was not yet explicitly implemented in the instructional materials, students needed assistance of the teacher or a fellow student for this step. Not all students may have asked for such assistance;
• students do not realise that the flow that must be defined is variable;
• students consider the model as a description of the situation that holds for just one moment;
• students do not use the notions of variable and formula because they have not yet sufficiently internalized these notions;
• students give in to their general tendency to consider only a number as a proper answer to a question;
• in spite of the computer tasks in the first two modules described in this paper, students do not realise that in computer programs variables must be defined by formulas.

At the end of the learning path, results suggest that students have started to understand and have internalised the role of direct relations in modelling. Classroom observations, screen recordings, and audio recordings from students working on the modelling tasks of Sound and of Force and Movement, show that direct relations are used for defining variables in models by almost all students (Section 6.9.3). The fact that all models from Force and Movement are based on formulas that are known to the students may have contributed to this apparent progress. In the test questions in Force and Movement, at the end of the learning path, almost all students use formulas for defining variables too. This does not prove yet that these students have fully internalised the notion of formula as required for modelling, because in these test questions they have been explicitly asked for formulas, so they did not have to decide whether a formula was required or not.

6.10.4 Answer to Research Question 4

To what extent can students construct simple direct relations?

The answer to Research Question 4 is not straightforward. In several test questions, students need to construct direct relations. In some of these questions students are not explicitly asked for a formula, but instead are asked what should be entered in a dialog window in the computer learning environment. Our results indicate that an important reason for not answering these questions correctly is that not all students realise that in such a case a formula is required (Section 6.10.3). Therefore, for analysing students’ construction abilities in itself, we concentrate on the questions in which students are explicitly asked to construct a formula.
The percentage of correct answers on test questions on formula construction of *Resistance and Conductivity* varies between 44% and 76% (Sections 6.7.4.2 and 6.7.4.3). This seems promising, since before that moment the students had never constructed formulas without teacher assistance. But alternative construction strategies, such as reasoning based on linguistic analogy, probably have had great influence.

Student results of the main pump test question from the next module, *The Vacuum Pump*, have considerably improved compared to the results of Van Buuren et al. (2012), but still the percentage of correct formulas is only 36%, even though students in this question only are asked to reconstruct a formula which they have used several times before. At the same time, after only a few exercises, two thirds of all students is able to construct a direct relation connecting three quantities (Section 6.8.4.5). Apparently, at this stage of their development, only a minor part of the students has problems with the construction of simple formulas per se, but a major part of the students has difficulties constructing a direct relation for a flow variable in a graphical model. A reason for this difference in performance is that the construction of the formula for the pump test question is the last step of a more extended process which must be understood completely. Varying quantities and another formula take part in this process and, as a separate step, an important factor (the pump factor) must be calculated first. In contrast, the formula in Section 6.8.4.5 can be generalised from just one calculation with three concrete placeholders.

At the end of the year, 50% of the students who did the final test of *Force and Movement* constructed a correct direct relation as definition for a flow variable in a new context.

The outcomes of the work of these lower secondary students are promising when compared to results from Westra (2008, p. 177) with pre-university upper secondary students (age: 16-17 years) in biology. Westra reports severe student difficulties in formalizing the relations between various components in a model. According to Westra, less than 35% of these upper secondary students were able to quantify causal relations.

An interesting question is, why we found so few reversed equation errors. One explanation is that the suggestion of Davis (1984), that students consider symbols as units or labels instead of true variables, may not hold for physics education, where students are used to distinguishing between units and variables from the beginning. The attention that is paid in our learning path to the notion of formula and to the notion of (symbolised) variable as representing a (possibly varying) quantity may be a second cause. Our approach, in which formulas generally are constructed by means of generalisation of calculations, may be a third reason. By offering this approach, we respond to the advice of Clement et al. (1981), to prevent students from solving such translation problems by means of trivial syntactic or other non-operative approaches.

We conclude that two thirds of the students following this learning path are able to construct simple direct relations involving concrete placeholders, but
a majority of the students still have difficulties with the construction of direct relations for flow variables. Students appear to be in need of strategies for formula construction. The numerical approaches that we have used in our learning path are useful when applied consciously. Students must become more aware of this. Also, learning to use proportionality arguments may be an effective approach. Such approaches are difficult, however, for contexts involving abstract, unfamiliar quantities that all can vary independently, such as the formulas from electricity in Section 6.7. In the history of physics, many formulas have not been found by reasoning, but have been derived from measurements. Therefore, an approach in which relations can be established by doing measurements seems appropriate for more abstract and complicated formulas. In order to be able to use such an approach, students must learn to derive mathematical relations from experimental results. Curve fitting can be a useful technique to do so. By using such an approach, students may simultaneously develop a better understanding of scientific methods. Designing such an approach is a valuable next step for the modelling learning path.

6.10.5 Answer to Research Question 5

To what extent can students interpret stock-flow diagrams as difference equations and vice versa?

In the initial phases of the third year of secondary education, at the end of The Vacuum Pump, only approximately 30% of all students can give the Δ-formula to a given simple new stock-flow diagram (Section 6.8.4.3). The most important error, made by 30% of all students, is the omission of any reference to the time step Δt.

At the end of this third year, more than two thirds of the students who did the final test of Force and Movement was able to construct a stock-flow diagram to a given Δ-formula, but this group of students may not be representative for all students. Therefore, a clear, satisfying answer to Research Question 5 cannot yet be given. But, even in this not representative group, 30% of the students add an icon for the time step when translating a Δ-formula into a stock-flow diagram (Section 6.9.4.1). We therefore conclude that stock-flow diagrams do not sufficiently communicate the role of the time step for many students. A second cause for this problem may be that for running graphical models, no explicit understanding of the relation between Δ-formulas and stock-flow diagrams is required. There is no clear need for students to understand this connection until stock-flow diagrams must be constructed based on Δ-formulas. As a result, the exact meaning of stock-flow diagrams can stay unclear for students for a long time. These problems with the time step must be dealt with before Research Question 5 can be answered satisfactory.
6.10.6 Answer to Research Question 6

To what extent can students construct simple difference equations?

Almost 40% of the students who did the final test of *Force and Movement*, can construct a Δ-formula without much practising beforehand. The quality of many incorrect answers is fairly good, and most incorrect answers are from ill-prepared students. But here also, there are difficulties with the time step in the formula. Almost two fifth of the students leave out this time step. Student errors also indicate that the role of initial values needs more attention (Section 6.9.4.2).

This group of third year students is probably not representative for all third year students. It can be expected that results of the entire group of third year students on this question would have been worse. Though, the results can be considered promising, because constructing difference equations is difficult even for upper secondary students (Verhoef et al., 2013), whereas the lower secondary students who have followed the learning path have hardly practised constructing Δ-formulas.

The problem of inclusion of the initial value into the Δ-formula may be related to the invisibility of the Δ-formula in the dialog window for the definition of a stock variable. If this Δ-formula would be visible together with the initial value, the special role of the initial value would be indicated more clearly. At the same time, the incorrect idea would be counteracted that stocks can be defined by means of constants. Some of the more subtle student errors, such as the replacement of Δ*B* by ‘*B* minus the initial value’ probably require a better understanding of the process of numerical integration.

6.10.7 General discussion of the partial learning path on variables and formulas

In this paper, we have demonstrated that students need proper conceptions of variable and formula for the approach to graphical modelling of our learning path. Such conceptions may be considered as rather abstract. Questions may be raised if such abstract conceptions cannot be avoided, by using another approach to graphical modelling or by starting with graphical modelling at a higher age. Another question is about the use of such abstract notions for students not applying for physics in upper secondary education. These questions are discussed in this section.

First, the question is addressed whether these abstract conceptions can be avoided if other approaches to graphical modelling are chosen. Often in introductory courses to graphical modelling, approaches are used that start with establishing causal relations between quantities. Also, regularly use is made of
the so-called stock-and-flow metaphor. In this approach, stock-flow diagrams are not explicitly connected to difference equations, but metaphorically to tanks and valves filled with water. In this way, understanding of difference equations is circumvented. For a detailed discussion of these approaches, we refer to Van Buuren et al. (2013b). Here, we just notice that most of the authors mentioned in Section 6.1 have used such approaches and report on difficulties that students have when designing or adapting graphical models (Bliss, 1994; Kurtz dos Santos & Ogborn, 1994; Sins et al., 2005; Lane, 2008; Westra, 2008; Ormel, 2010).

The second question is whether such problems with formulas would also have arisen if we had postponed the start of the modelling learning path. In that case, we maybe could have used formulas in graphical models right from the start. But results of the questionnaire mentioned in Section 6.1 indicate that similar misconceptions with respect to formulas and numbers would have evolved as well. Besides, other distorting conceptions would have got time to develop. An example is the improper use of closed mathematical solutions by students in modelling activities, such as the use of \( v = a \cdot t \) for defining a flow variable \( v \) (velocity), even in case the acceleration \( a \) is not constant, and the erroneous use of \( x = \frac{1}{2} a \cdot t^2 \) for the initial value of the stock variable \( x \) (see also Van Buuren et al., 2010). Finally, we note that part of the authors referred to in Section 6.1 report on difficulties with respect to design and adaptation of models of upper secondary students (Sins et al., 2005; Westra, 2008; Ormel, 2010). Apparently, starting with modelling at a higher age has not prevented such difficulties.

As a more general reply to these first two questions, we note that an increased awareness of students of the varying character of quantities is not only important for modelling, but is useful in more traditional education as well. In physics courses, students often have to find solutions that hold for only one moment in situations in which, in reality, quantities are changing permanently. Examples are the calculations of momentary velocities in kinematics and of problems involving momentary two dimensional forces in dynamics. The notion of variable as varying quantity is useful in clarifying the difference between momentary values and constants. Another example is the difference in vibration theory between time (a variable) and period (a constant). In mathematics education, Tall and Thomas (1991) have shown that a proper notion of variable contributes significantly to the understanding of higher order concepts in algebra, and that an initial loss of manipulative skills as a result of the increased attention to the notion of variable, is more than made up at a later stage.

The third question is why we should burden lower secondary students who do not intend to apply for physics in upper secondary education with such abstract concepts. Our reply would be that the lower secondary students who have followed this learning path, have developed a valuable base of orientation with respect to important concepts for modelling. It can serve as a good foundation for modelling activities in upper secondary education in other disciplines, such as mathematics, biology, economy, and chemistry, as well. Motivation for doing physics of students who do not apply for physics in the next grade often
decreases towards the end of lower secondary education. These students keep more motivation for the lower secondary physics curriculum when they understand that modelling competencies acquired in physics lessons are useful in other disciplines as well.

6.10.8 Recommendations for further design and research

From results presented in Section 6.10.3 it follows that the last steps of the learning trajectory on the conceptions of variable and formula for modelling needs some refinement. More explicit attention must be paid in the instructional materials to the notion that a variable in a computer model needs to be defined by means of a formula.

In order to find out to what extent the first models of the learning path, in which no formulas are used explicitly, have induced the misconception that in computer models variables can be defined by numbers only, these first models should be adapted to make the role of the difference equation and its connection to the stock-flow diagram more explicit. Such an approach may also be effective in preventing the problems with the time step and the initial value reported in Sections 6.10.5 and 6.10.6. One approach can be to use the option incorporated in Coach to switch between the mode for graphical modelling and the mode for text modelling. This may clarify the relation between the graphical diagram and the integration in a more explicit way than we have used so far. Another solution may be the adaptation of the graphical diagram, so that it communicates the relation with the formula more clearly. This may be done by presenting the Δ-formula in the computer learning environment together with the stock-flow diagram, instead of hiding this formula. This possibility could be added to the software as an option, for educational purposes. As soon as students have come to understand this relation, the Δ-formulas can be left out.

In Section 6.10.4, we stated that students appear to be in need of strategies for formula construction. We propose adaptation of the design of the instructional materials, in a way that students are stimulated to apply construction strategies more consciously. For more abstract formulas, appearing at later stages of physics education, an approach must be developed in which formulas can be derived from measurements.

References


162


Schaap, S., Vos, P., Ellermeijer, T., & Goedhart, M. (2011). De vertaalslag van een situatie naar een wiskundige formule; een studie naar vraagstellingen en


